Numerical Approximations for Stochastic Systems With Delays in the State and Control

Harold J. Kushner*
Applied Mathematics Dept, Brown University
Providence RI 02912
hjk@dam.brown.edu

December 26, 2005

Running head: Numerical Methods for Stochastic Delay Systems

Abstract

The Markov chain approximation numerical methods are widely used to compute optimal value functions and controls for stochastic as well as deterministic systems. We extend them to controlled general nonlinear delayed reflected diffusion models. The path, control, and reflection terms can all be delayed. Previous work developed numerical approximations and convergence theorems. But when the control and reflection terms are delayed those and all other current algorithms normally lead to impossible demands on memory. An alternative "dual" approach was proposed by Kwong and Vintner for the linear deterministic system with a quadratic cost function. We extend the approach to the general nonlinear stochastic system, develop the Markov chain approximations and numerical algorithms, and prove the convergence theorems. The approach reduces the memory requirement significantly. For the no-delay case, the method covers virtually all models of current interest. The method is robust and the approximations have physical interpretations as control problems closely related to the original one. These advantages carry over to the delay problem.

Keywords: Optimal stochastic control, numerical methods, delay stochastic equations, numerical methods for delayed controlled diffusions, Markov chain approximation method.
2000 Mathematics Subject classifications: AMS numbers: 93E20, 65C30, 65Q05, 65H35, 34K28, 34K50, 93E03

*This work was partially supported by NSF grant DMS-0506928 and ARO contract W911NF-05-10928

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **26 DEC 2005**	2. REPORT TYPE	3. DATES COVERED **00-12-2005 to 00-12-2005**

4. TITLE AND SUBTITLE **Numerical Approximations for Stochastic Systems With Delays in the State and Control**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Brown University,Division of Applied Mathematics,182 George Street,Providence,RI,02912**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**36**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

1 Introduction

The paper [6] extended the numerical methods of [10], known as the Markov chain approximation methods, to controlled delayed diffusion models, to numerically obtain the optimal costs and controls. The state of the model, as needed for the numerical procedure, consists of the segment of the path over the delay interval and of the control path as well (if the control is also delayed). Delayed reflection terms were not dealt with. Convergence theorems were proved and "numerically efficient" representations of the state data were developed that reduced the memory requirements to manageable size for low-dimensional problems, if the path only were delayed. If the control and/or reflection terms are also delayed, then the memory requirements with any current method is prohibitive.

In this paper we will take an alternative approach that greatly reduces the memory requirements for general nonlinear stochastic problems where the control and reflection terms, as well as the path variables, are delayed. The approach was suggested by the work in [13] which dealt only with the linear deterministic system with a quadratic cost function, and the development depended heavily on the linear structure. But the idea can be extended to the problem of concern here and has numerous advantages.[1] With this method, the delay equation is replaced by a type of stochastic wave equation with no delays, and its numerical solution yields the optimal costs and controls for the original model. With appropriate numerical algortihms the memory requirements are much reduced over more direct methods.

Delayed reflection terms occur frequently in applications to communications systems, where one of the reflection terms corresponds to buffer overflow. This data is then communicated to the controller via a transportation delay. In fact such problems have been a major motivation for this work and an example is given in the next section. There is an large literature on the delayed linear system, quadratic cost, and white Gaussian noise case [2, 7, 9, 11, 13]. As for the no-delay case, the analysis is essentially the same with and without driving noise. The problem reduces to the study of an abstract Ricatti equation. But little has been available for the general nonlinear stochastic problem.

The basic idea of the numerical method is to first approximate the control problem by a control problem with a suitable approximating Markov chain model, then solve the Bellman equation for the approximation, and then prove the convergence of the optimal costs to that for the original problem as the approximation parameter goes to zero. The method is robust and the approximations have physical interpretations as control problems closely related to the original one.

Models for many physical problems have reflecting boundaries. They occur naturally in models arising in queueing/communications systems [8], where the state space is often bounded owing to the finiteness of buffers and nonnegativity of their content, and the internal routing determines the reflection directions on

[1] The author would like to thank Kasi Itô for bringing the paper to his attention.

the boundary. See the example in Section 2. Numerical analysis for dynamical models is usually done in a bounded region. A common way of bounding a state space is to impose a reflecting boundary, if one does not exist already. One selects the region so that the boundary plays a minor role. Alternatively, one can stop the process on exit from a given bounded set. We work with the reflecting boundary case, since it is of considerable importance. The modifications that are required when the boundary is absorbing are minor.

The model and assumptions are in Section 2. The assumptions on the reflection directions are those in [10] and are standard for the reflected diffusion model (also known as the Skorokhod problem). Section 3 is concerned with a representation of the solution in terms of a type of stochastic wave equation without delay terms. This is an extension to the general nonlinear stochastic system of the idea in [13] for the deterministic linear problem. The reference contains a history of the idea, still for the linear deterministic system. With this representation, the delays are eliminated, but one must solve a PDE. It is shown that the representation is equivalent to the original problem in that any solution to one yields a solution to the other. Owing to the special form of the PDE, there are straightforward ways of getting numerical approximations.

To prepare ourselves for what will be required for the numerical approximations, a discrete-time approximation is developed in Section 4. This will suggest the correct scaling and illustrate the type of algebraic manipulations that are required. It might also be a convenient way of simulating the original system. The Markov chain approximation method is reviewed in Section 5, starting with the simpler no-delay case. The types of continuous-time interpolations that will be of interest are outlined and the asymptotic equivalence of their scalings is proved. We try to set the problem up so that the methods and results of [10] can be used without excessive duplication of details and the proofs in [10] can be appealed to and used where possible. The form of the transition probabilities for the approximating Markov chain are given. The fundamental assumption required for the convergence of the numerical procedure is the so-called "local consistency condition" [10]. This says little more than that the conditional mean change (resp, variance) in the state of the approximating chain is proportional to the drift (resp, covariance) of the diffusion, modulo small errors. This would seem to be a minimal condition. In general, it need not hold everywhere (see, e.g., [10, Section 5.5]).

The proofs of convergence in [10] are purely probabilistic, being based on weak convergence methods. The idea is to interpolate the chain to a continuous-time process in a suitable manner, and then show that the interpolated processes converge to an optimal diffusion as the approximating parameter goes to zero. The Skorohod topology is used on the path spaces. The criterion for tightness for vector-valued processs is [10, Theorem 2.1, Chapter 9].

Section 6 contains the main development of the approximation for the delay case. Motivated by the ideas in Sections 4 and 5, it is shown how to construct the approximating chains for the representation introduced in Section 3. The method is very close to that used for the no-delay case. Representations of the resulting process are developed that facilitate proving that the optimal costs for the approximating chain converge to that for the original diffusion. The development is designed so that the convergence proofs in [10] can be appealed to. In Section 7, the size of the state space that is needed for the solution of the Bellman equation is discussed, and it is seen that the approach does moderate the requirements considerably. Although the form of the algorithm is motivated by those used for the no-delay problem, it is more complicated. But then the delay problem is substantially more complicated and the proposed algorithm seems considerably better in terms of memory requirement than any current alternative if the control and/or reflection terms are delayed.

The algorithms can be used to obtain optimal value functions or controls. Controls for delay systems will usually be too complicated for direct practical use. But the values provide benchmarks. One would normally try to approximate the optimal control by realizable forms. For example, one might try to approximate the control for the example in Section 2 by a rate dependent threshold on the buffer size. The algorithms can be used for numerical exploration. For example, to explore the effect of changing delay, under optimality conditions. Generally, a cost function is a compromise between competing criteria, and the values of the individual components of the cost as well as the total value are important. By varying the weights or the form of the components, the algorithms give the tradeoffs between components under the best conditions, namely under the optimal controls. All of this is very useful information for the designer.

2 The Model and Assumptions

A motivating example. Due to the finite speed of electromagnetic signaling, delays are a common and crucial part of many telecommunications systems. One important example is the AIMD (additive increase multiplicative decrease) model that arises in (FTP) control of internet traffic over long distances. The following model, from [1], is a good example.

$$dx_1(t) = c_1 dt - \kappa_0 dy_{22}(t-\tau) + \kappa_1 u_1(x_2(t-\tau), x_1(t-\tau))dt + dz_1(t),$$

$$x_2(t) - x_2(0) = \int_0^t [x_1(s) - b]\, ds + w(t) + z_2(t). \tag{2.1}$$

Here c_1 and b are constants, $w(\cdot)$ is a real-valued Wiener process, $x_2(\cdot)$ represents a scaled buffer content and $x_1(\cdot)$ is a scaled and centered rate of transmission. The buffer and source are separated by a channel with a transmission delay. The buffer is finite and $0 \le x_2(t) \le B$. The reflection term $z_2(\cdot)$ serves to keep $x_2(\cdot)$ in the desired range. We write $z_2(\cdot) = y_{21}(\cdot) - y_{22}(\cdot)$ where the nondecreasing reflection component $y_{22}(\cdot)$ represents buffer overflows (lost packets) and can increase only when $x_2(t) = B$. The nondecreasing component $y_{21}(\cdot)$ keeps the buffer content nonnegative. In this model, overflow packets are not acknowledged. Hence after the communication delay τ, the lack of acknowledgment causes an automatic decrease in the transmission rate $x_1(\cdot)$. The $u_1(\cdot)$ is a control that allows the user to adjust its transmission rate as a function of the delayed data. The scaled rate variable $x_1(\cdot)$ might not be physically bounded. But for the purposes of numerical approximations, it is usually confined to an appropriate set and kept in that set by means of boundary reflection $z_1(\cdot)$. Note that the delayed reflection term $y_{22}(\cdot)$ is an essential component of the model.

The model. The general model has the form of a controlled reflected diffusion process with possible delays in the state, control, and reflection term. Let $I\!R^r$ denote Euclidean r-space. The r-dimensional state process $x(\cdot)$ will be confined to a convex polyhedron $G \in I\!R^r$, with a nonempty interior, by means of the boundary reflection process $z(\cdot)$, as in [6]. The conditions on the reflection directions on the boundary are spelled out below. Let x_i denote the ith component of a vector x. Note that what is usually penalized are buffer overflows, and on the boundaries associated with overflows, the reflection directions are usually inward normals to the boundary.

Assumptions on the state space G. Assumptions (A2.1) and (A2.2) are the ones used in [10] (see Section 5.7 of that reference), and are standard in the treatment of general reflecting diffusions [3, 4], [8, Section 3.5].

A2.1. *The state space G is the intersection of a finite number of closed half spaces in Euclidean r-space $I\!R^r$ and is the closure of its interior (i.e., it is a closed convex polyhedron with an interior and planar sides). Let ∂G_i, $i = 1, \ldots,$ denote the faces of G, and n_i the interior normal to ∂G_i. Interior to ∂G_i, the reflection direction is denoted by the unit vector d_i, and $\langle d_i, n_i \rangle > 0$ for each i. The possible reflection directions at points on the intersections of the ∂G_i are in the convex hull of the directions on the adjoining faces. Let $d(x)$ denote the set of reflection directions at the point $x \in \partial G$, whether it is a singleton or not. No more than r constraints are active at any boundary point.*

A2.2. *For each $x \in \partial G$, define the index set $I(x) = \{i : x \in \partial G_i\}$. Suppose that $x \in \partial G$ lies in the intersection of more than one boundary; that is, $I(x)$ has the form $I(x) = \{i_1, \ldots, i_k\}$ for some $k > 1$. Let $N(x)$ denote the convex hull of the interior normals $n_{i_1}, \ldots, n_{i_k}$ to $\partial G_{i_1}, \ldots, \partial G_{i_k}$, resp., at x. Then there is some vector $v \in N(x)$ such that $\gamma' v > 0$ for all $\gamma \in d(x)$.*

There is a neighborhood $N(\partial G)$ and an extension of $d(\cdot)$ to $\overline{N(\partial G)}$ that is upper semicontinuous in the following sense: For each $\epsilon > 0$, there is $\rho > 0$ that goes to zero as $\epsilon \to 0$ and such that if $x \in N(\partial G) - \partial G$ and distance$(x, \partial G) \le \rho$, then $d(x)$ is in the convex hull of the directions $\{d(v); v \in \partial G, \text{distance}(x, v) \le \epsilon\}$.

Our general model takes the form, where we define $d\mu(\theta) = \mu(\theta) - \mu(\theta - d\theta)$,

$$\begin{aligned} dx(t) = c(x(t), u(t))dt + dt \int_{-\tau}^{0} b(x(t+\theta), u(t+\theta), \theta)d\theta \\ +dt \int_{-\tau}^{0} g(x(t+\theta), u(t+\theta), \theta)d\mu(\theta) + \sigma(x(t))dw(t) + dz(t) \\ +dt \int_{\theta=-\tau}^{0} p(\theta)d_\theta y(t+\theta). \end{aligned} \tag{2.2}$$

Here $w(\cdot)$ is a standard and possibly vector-valued Wiener process, $z(\cdot)$ is the reflection term and $\tau > 0$ is the non-random maximum delay. We can represent $z(\cdot)$ as $z(t) = \sum_i d_i y_i(t)$, where $y_i(\cdot)$ denotes the nondecreasing process that can increase only when $x(t)$ is on the ith face of G. The last integral in (2.2) is with respect to $d\theta$ in the sense that

$$p(\theta)d_\theta y(t+\theta) = p(\theta)\left[y(t+\theta+d\theta) - y(t+\theta)\right].$$

The initial condition is the pair $\hat{x} = \{x(s), -\tau \leq s \leq 0\}$, $\hat{u} = \{u(s), -\tau \leq s \leq 0\}$. One could incorporate the term containing $b(\cdot)$ into the term containing $g(\cdot)$. But the idea is that the first term represents distributed delays, while the latter can represent "point" delays, " so we prefer to use both terms. See [8, 10] for more detail on controlled reflected diffusions.

Note that in the example (2.1) the delayed reflection term is $y_{22}(\cdot)$, a component of $z_2(\cdot)$, and it appears only in the equation for $x_1(\cdot)$. Since $x_1(\cdot)$ is bounded, the processes $x_2(\cdot), z_2(\cdot)$ are well defined and hence the term $y_{22}(\cdot)$ in the equation for $x_1(\cdot)$ is well defined; it is continuous and nondecreasing. In order to assure that the reflection and solution for (2.2) are well defined we need to assure that the delayed reflections and non-delayed reflections are "separated." This consideration and the structure of the motivating example lead to the following assumption for the $p(\cdot)$ in (2.2).

A2.3. *There is $\tau_0 \in (0, \tau]$ such that $p(\theta) = 0$ for $\theta \geq -\tau_0$.*

A2.4. *The control takes values in a compact set U and is measurable (as a function of (ω, t)) and is nonanticipative with respect to $w(\cdot)$. The functions $b(\cdot), c(\cdot), p(\cdot), g(\cdot)$ are bounded and continuous. All functions of θ have value zero for $\theta < -\tau$ and $\theta > 0$. The function $\sigma(\cdot)$ is bounded and continuous and $\mu(\cdot)$ is a finite measure on $[-\tau, 0]$ with $\mu([-t, 0]) \to 0$ as $t \to 0$. We suppose that $z(t) = 0, t \leq 0$.*

A2.5. *There is a unique weak sense solution to (2.2) for each initial condition and admissible control.*

The reflected diffusion model (2.2) is known as the Skorokhod problem. For a detailed discussion of the Skorohod problem and the assumptions (A2.1) and (A2.2), see [8, Chapter 3], [3, 4]. Let $|z|(t)$ denote the variation of $z(\cdot)$ on the interval $[0,t]$. By a solution to (2.2) we mean the following. The $z(\cdot)$ is the reflection process and satisfies the following conditions: $|z|(t) < \infty$ with probability one (w.p.1) for all t, and there is a measurable function $\gamma(\cdot)$ with $\gamma(t) \in d(x(t))$ w.p.1 such that $z(t) = \int_0^t \gamma(s) d|z|(s)$. This says only that the reflection process can change only when $x(t)$ is on the boundary, and the increments are in a correct reflection direction.

Comments on the assumptions. One can always construct the extension in (A2.2). Under (A2.1)–(A2.2), the choice of the reflection direction on the corners and edges of G has no effect on the process. To see that (A2.1) is natural in applications note the following. If the state space is being bounded for purely numerical reasons, then the reflections are introduced only to give a compact set G, which should be large enough so that the effects on the solution in the region of main interest are small. A common choice is a hyperrectangle with interior normal reflection directions. The condition (A2.2) implies (see [3, 8]), the so-called "completely-S" condition, the fundamental boundary condition for the modelling of stochastic networks, [5, 8, 12], and which is used to ensure that $z(\cdot)$ has bounded variation w.p.1. Various extensions are possible. A jump term can be added with no additional problems, provided that it does not involve a delay. The $x(\cdot)$ could also be delayed in the function $\sigma(\cdot)$. A delayed Wiener process term such as

$$dt \int_0^t \sigma^1(x(t+\theta),\theta) d_\theta w(t+\theta)$$

can be added.

Comment on the delayed reflection term in (2.2). Consider a one-dimensional problem. Let $p(\theta) = 1/\delta$ on the interval $[-\Delta-\delta, -\Delta]$, $\Delta > 0, \delta > 0$, with value zero elsewhere. Then, with $y(t) = 0, t \leq 0$, we have

$$\int_0^t ds \int_{-\tau}^0 p(\theta) d_\theta y(s+\theta) = \frac{1}{\delta} \int_{t-\Delta-\delta}^{t-\Delta} y(s) ds \approx y(t-\Delta)$$

for small δ. In this way, point delays as well as distributed delays can be approximated.

Relaxed controls. For purposes of proving approximation and limit theorems, it is usual and very convenient to work in terms of relaxed controls. Recall the definition of a relaxed control $m(\cdot)$ [10]. It is a measure on the Borel sets of $U \times [0,\infty)$, with $m(A \times [0,\cdot])$ being measurable and nonanticipative with respect to $w(\cdot)$ for each Borel $A \in U$, and satisfying $m(U \times [0,t]) = t$. Write $m(A,t) = m(A \times [0,t])$. The left-hand derivative[2] $m'(d\alpha, t) = \lim_{0<\delta\to 0}[m(d\alpha,t) - m(d\alpha, t-\delta)]/\delta$ is defined for almost all (ω, t). By the definitions, $m(d\alpha\, ds) = m'(d\alpha, s)ds$. For $0 \leq v \leq \tau$, we write $m(d\alpha, ds - v)$ for

[2] In [10] m_t was used to denote the derivative. But this notation would be confusing in the context of the notation required to represent the various delays in this paper.

$m(d\alpha, s - v) - m(d\alpha, s - ds - v)$. The weak topology is used on the relaxed controls. Thus $m^n(\cdot)$ converges to $m(\cdot)$ if and only if $\int\int \phi(\alpha, s) m^n(d\alpha\, ds) \to \int\int \phi(\alpha, s) m(d\alpha\, ds)$ for all continuous functions $\phi(\cdot)$ with compact support. With this topology, the space of relaxed controls is compact. An ordinary control $u(\cdot)$ can be written as the relaxed control $m(\cdot)$ defined by its derivative $m'(A, t) = I_{\{u(t) \in A\}}$, where I_K is the indicator function of the set K. Then $m(A, t)$ is the amount of time that the control takes values in the set A by time t. The controls for the numerical procedure will be ordinary controls, but those for the limit might be relaxed. For the initial condition on $[-\tau, 0]$ we use an ordinary control. The relaxed control form of (2.2) is

$$\begin{aligned} dx(t) = dt \int_U c(x(t), \alpha) m'(d\alpha, t) + dt \int_U \int_{-\tau}^0 b(x(t+\theta), \alpha, \theta) m'(d\alpha, t+\theta) d\theta \\ + dt \int_{-\tau}^0 \int_U g(x(t+\theta), \alpha, \theta) m'(d\alpha, t+\theta) d\mu(\theta) + \sigma(x(t)) dw(t) \\ + dz(t) + dt \int_{\theta=-\tau}^0 p(\theta) d_\theta y(t+\theta). \end{aligned} \tag{2.3}$$

The following theorem is [10, Theorem 1.1, Chapter 11]. The last assertion is proved by working recursively on intervals (see (A2.3) for the definition of τ_0) $(0, \tau_0], (\tau_0, 2\tau_0], , \ldots,$ until the interval $(0, T]$ is covered.

Lemma 2.1. *Assume* (A2.1)–(A2.2). *Let $f(\cdot)$ and $\sigma(\cdot)$ be measurable and nonanticipative processes of the appropriate dimension, and bounded in absolute value by some constant $K < \infty$. Define*

$$dX(t) = f(t)dt + \sigma(t)dw(t) + dZ(t), \quad X(0) \in G,$$

where $Z(\cdot)$ is the reflection term. Let $|Z|(t)$ denote the variation of $Z(\cdot)$ on the interval $[0, t]$. Then

$$\lim_{T \to 0} \sup_{X(0), f, \sigma} E|Z|^2(T) = 0. \tag{2.4}$$

For each $T < \infty$,

$$\sup_{X(0), f, \sigma} E|Z|^2(T) < \infty. \tag{2.5}$$

Let $Y_i(\cdot)$ denote the component of the reflection process that is due to reflection on the ith face, with corners and edges assigned in any way at all to the adjacent faces. Assume the condition (A2.3) *on $p(\cdot)$, and redefine $X(\cdot)$ by*

$$dX(t) = f(t)dt + \sigma(t)dw(t) + dt \int_{\theta=-\tau}^0 p(\theta) d_\theta Y(t+\theta) + dZ(t), \quad X(0) \in G.$$

Then (2.4) *and* (2.5) *continue to hold.*

With $p(\cdot)$ omitted, the bounds in (2.4) and (2.5) depend on $E \sup_{t \le T} |X(0) + \int_0^t f(s)ds + \int_0^t \sigma(s)dw(s)|^2$, and the appropriate "recursive" adjustments are made when $p(\cdot) \neq 0$.

An extension. Suppose that the system evolves as $dX(t) = f(t)dt+\sigma(t)dw(t)+ dt\int_{\theta=-\tau}^{0} p(\theta)d_\theta Y(t+\theta)$ on the intervals $[n\Delta, n\Delta+\Delta)$, and the reflection comes in at times $n\Delta$ if $X(n\Delta-) \notin G$. Then the lemma holds if $\lim_{T,\Delta\to 0}$ replaces $\lim_{T\to 0}$ in (2.4). The proof is similar to that in the reference.

A discounted cost function. Let $\hat{x}$ and $\hat{u}$ denote the canonical value of the path and control segments, resp., on $[-\tau, 0]$. For $\beta > 0$, some vector q, and control process $u(\cdot)$ on $[0,\infty)$, the cost function is

$$W(\hat{x},\hat{u},u) = E_{\hat{x},\hat{u}}^{u} \int_0^\infty e^{-\beta t}\left[k(x(t),u(t))dt + q'dy\right], \tag{2.6}$$

with the analogous form for a relaxed control. The function $k(\cdot)$ is bounded, continuous, and real-valued. If two adjacent faces of G have the same reflection direction, then the associated components of the vector q must be the same. By Lemma 2.1, the reflection term component of the cost is well defined. The $E_{\hat{x},\hat{u}}^{u}$ denotes the expectation given the initial condition and that control $u(\cdot)$ is used. Let $V(\hat{x},\hat{u})$ denote the infimum of the costs, over all controls.

Existence of an optimal control. The existence of an optimal relaxed control was shown in [6, Theorem 2.1] for a one-dimensional model without a delayed reflection term. The proof in the case of concern here is similar. One takes minimizing sequences of controls and then a weakly convergent subsequence of the (path, relaxed control, Wiener process, reflection term) and shows that the limit satisfies (2.3) and that the minimizing sequence of costs converges to the cost for the limit processes. Furthermore ([6, Theorem 2.3]) the infimum of the cost over the relaxed controls is equal to the infimum of the costs over ordinary controls.

3 A Useful Representation of $x(\cdot)$

For $-\tau < \theta \le 0$, define processes $\chi^0(\cdot)$ and $\chi^1(\cdot)$ by

$$d\chi^0(t) = \chi^1(t,0)dt + c(\chi^0(t),u(t))dt + \sigma(\chi^0(t))dw(t) + dz^0(t), \tag{3.1}$$

$$\begin{aligned} d_t\chi^1(t,\theta) = &-d_\theta\chi^1(t,\theta) + b(\chi^0(t),u(t),\theta)dt \\ &+g(\chi^0(t),u(t),\theta)\left[\mu(\theta)-\mu(\theta-dt)\right] + p(\theta)dy^0(t). \end{aligned} \tag{3.2}$$

The reflection term $z^0(\cdot)$ will be for the process $\chi^0(\cdot)$, which takes values in the constraint set G and is subject to the boundary conditions (A2.1) and (A2.2). The interpretation of the stochastic partial differential equation (3.2), as well as of the relaxed control form (3.2r) below, is given by (3.5), (3.6) below, which defines the solution. Theorem 3.1 shows that $\chi^0(\cdot) = x(\cdot)$. If there is no delayed reflection term, then the values of $\chi^1(t,\theta)$ will be seen to be bounded. If the solution is not bounded, then "numerical" bounds will have to be added, and we return to this point in Section 6. The relaxed control forms of (3.1) and (3.2) are

$$d\chi^0(t) = \chi^1(t,0)dt + \int_U c(\chi^0(t),\alpha)m'(d\alpha,t)dt + \sigma(\chi^0(t))dw(t) + dz^0(t), \tag{3.1r}$$

$$\begin{aligned} d_t\chi^1(t,\theta) &= -d_\theta\chi^1(t,\theta) + dt\int_U b(\chi^0(t),\alpha,\theta)m'(d\alpha,t) \\ &\quad + \int_U g(\chi^0(t),\alpha,\theta)m'(d\alpha,t)\left[\mu(\theta)-\mu(\theta-dt)\right] + p(\theta)dy^0(t). \end{aligned} \tag{3.2r}$$

These processes will be the basis of the development of the numerical method. The linear and deterministic form of (3.1), (3.2) were used in [13] to represent the linear and deterministic analog of (2.2), but without the analog of the terms involving $\mu(\cdot)$ or the reflection.

The initial conditions for (3.1) and (3.2) are $\chi^0(0) = x(0)$, $z^0(s) = z^1(s,\theta) = 0$ for $s \le 0$, and

$$\begin{aligned} \chi^1(0,\theta) &= \int_{-\tau}^{\theta} b(x(\gamma-\theta),u(\gamma-\theta,\gamma)d\gamma \\ &\quad + \int_{-\tau}^{\theta} g(x(\gamma-\theta),u(\gamma-\theta),\gamma)d\mu(\gamma) + \int_{-\tau}^{\theta} p(\gamma)d_\gamma y^0(\gamma-\theta). \end{aligned} \tag{3.3}$$

The boundary condition is $\chi^1(t,-\tau) = 0$.

Note on dimension and size of the system state. The dimension of $\chi^1(\cdot)$ is equal to the number of components of $x(\cdot)$ whose dynamical terms have delays. Thus the method would currently be impractical if the dynamics of more than one component of x contained delays. For the components $x_i(\cdot)$ whose dynamical terms do not have delays, simply define $\chi_i^0(\cdot) = x_i(\cdot)$, $\chi_i^1(\cdot) = 0$. The dimension of $\chi^1(\cdot)$ does not depend on the number of controls. Suppose that delayed values of components $x_i(\cdot), i = 1,\ldots,r_1$, $u_i(\cdot), i = 1,\ldots,r_2$, and $y_i(\cdot), i = 1,\ldots,r_3$, are required. For the original problem, the full system state consists of the initial condition $x(0)$ and the memory segments of the $x_i(\cdot), i = 1,\ldots,r_1$, $u_i(\cdot), i = 1,\ldots,r_2$, and $y_i(\cdot), i = 1,\ldots,r_3$, where the $u_i(\cdot)$ have no particular regularity properties and are difficult to approximate efficiently. The full system or memory state for (3.1), (3.2), is just $x(0)$ and the current values $\chi^1(t,\theta)$, $-\tau \le \theta \le 0$.

A semigroup representation of (3.1), (3.2). The part $d_t\chi^1(t,\theta) = -d_\theta\chi^1(t,\theta)$ of (3.2) is a type of wave equation and its semigroup will play a major role. Following [13], define the semigroup $\Phi(\cdot)$ (where $-\tau \le \theta \le 0$) by

$$(\Phi(t)f(\cdot))(\theta) = \begin{cases} f(\theta-t), & -\tau \le \theta - t \le 0, \\ 0, & \text{otherwise}. \end{cases} \tag{3.4}$$

$\Phi(\cdot)$ will only act on functions of θ.

The construction of the numerical approximations will use the dynamical representation (3.1), (3.2) as a heuristic guide, but the solution to (3.1), (3.2) is always interpreted in the "variation of constants" form

$$d\chi^0(t) = \chi^1(t,0)dt + c(\chi^0(t), u(t))dt + \sigma(\chi^0(t))dw(t) + dz^0(t), \tag{3.5}$$

$$\begin{aligned}\chi^1(t,\theta) = \Phi(t)\chi^1(0,\theta) + \int_0^t \Phi(t-s)\left[b(\chi^0(s), u(s), \theta)ds + p(\theta)dy^0(s)\right] \\ + \int_0^t \Phi(t-s)g(\chi^0(s), u(s), \theta)\left[\mu(\theta) - \mu(\theta - ds)\right].\end{aligned} \tag{3.6}$$

The integral involving $\mu(\cdot)$ is well defined, since the integration is to be done after the operation by $\Phi(t-s)$ and we can write

$$\begin{aligned}&\int_0^t \Phi(t-s)g(\chi^0(s), u(s), \theta)\left[\mu(\theta) - \mu(\theta - ds)\right] \\ &\quad = \int_0^t g(\chi^0(s), u(s), \theta - t + s)I_{\{-\tau\le\theta-t+s\le 0\}}\left[\mu(\theta - t + s) - \mu(\theta - t + s - ds)\right] \\ &\quad = \int_{\max\{\theta-t,-\tau\}}^{\theta} g(\chi^0(\gamma + t - \theta), u(\gamma + t - \theta), \gamma)d\mu(\gamma).\end{aligned} \tag{3.7}$$

For the relaxed control form of (3.6), use the forms

$$\int_0^t \Phi(t-s)\int_U b(\chi^0(s), \alpha, \theta)m'(d\alpha, s)ds,$$

$$\int_0^t \int_U \Phi(t-s)g(\chi^0(s), \alpha, \theta)m'(da, s)\left[\mu(\theta) - \mu(\theta - ds)\right],$$

and in (3.5) use $dt\int_U c(\chi^0(t), \alpha)m'(d\alpha, t)$. The cost function is (2.6) with $\chi^0(t)$ replacing $x(\cdot)$.

The following theorem is a nonlinear and stochastic version of the linear and deterministic result in [13].

Theorem 3.1. *Assume* (A2.1)–(A2.5). *Then* (3.1) *and* (3.2) *have the weak sense unique solution*

$$\chi^0(\cdot) = x(\cdot), \tag{3.8}$$

$$\begin{aligned}\chi^1(t,\theta) = \int_{-\tau}^{\theta} b(\chi^0(t+\gamma-\theta), u(t+\gamma-\theta), \gamma)d\gamma + \int_{-\tau}^{\theta} p(\gamma)d_\gamma y^0(t+\gamma-\theta) \\ + \int_{-\tau}^{\theta} g(\chi^0(t+\gamma-\theta), u(t+\gamma-\theta), \gamma)d\mu(\gamma).\end{aligned} \tag{3.9}$$

The analogous result holds for the relaxed control form, where we use

$$\int_{-\tau}^{\theta}\int_U b(\chi^0(t+\gamma-\theta), \alpha, \gamma)m'(d\alpha, t+\gamma-\theta)d\gamma,$$

$$\int_{-\tau}^{\theta}\int_U g(\chi^0(t+\gamma-\theta),\alpha,\gamma)m'(d\alpha,t+\gamma-\theta)d\mu(\gamma)$$

in place of the first and third terms on the right side of (3.9).

Comment on uniqueness. We assumed that (2.2) and (2.3) have unique weak-sense solutions for any admissible control. The proof shows that any solution to (3.5), (3.6) has the form (3.8), (3.9). Given $x(\cdot)$ satisfying (2.2) or (2.3), replace $\chi^0(\cdot)$ in (3.9) by $x(\cdot)$. Setting $\theta = 0$ and substituting (3.9) into (3.5) with $\chi^0(\cdot) = x(\cdot)$ yields that (3.5) is just (2.2). Conversely, given a solution to (3.5), (3.6), where $\chi^1(\cdot)$ is given by (3.9), we have that $\chi^0(\cdot)$ solves (2.2). Hence, the solution to (3.5), (3.6) is also weak-sense unique.

Proof. For simplicity in the notation, work with ordinary rather than relaxed controls. The development for the relaxed control form is analogous.

From the comments above concerning uniqueness, we need only show that any measurable and non-anticipative solution $(\chi^0(\cdot), \chi^1(\cdot))$ satisfying (3.5), (3.6) implies (3.9). Consider the representation (3.6). The component due to the initial condition in (3.3) is

$$\begin{aligned}\Phi(t)\chi^1(0,\theta) &= \int_{-\tau}^{\theta-t} b(\chi^0(\gamma-\theta+t),u(\gamma-\theta+t),\gamma)I_{\{-\tau\le\theta-t\le 0\}}d\gamma \\ &\quad + \int_{-\tau}^{\theta-t} g(\chi^0(\gamma-\theta+t),u(\gamma-\theta+t),\gamma)I_{\{-\tau\le\theta-t\le 0\}}d\mu(\gamma) \\ &\quad + \int_{-\tau}^{\theta-t} p(\gamma)d_\gamma y^0(\gamma-\theta+t)\end{aligned}$$

$$\begin{aligned}&= \int_{-\tau}^{\max\{\theta-t,-\tau\}} b(\chi^0(\gamma-\theta+t),u(\gamma-\theta+t),\gamma)d\gamma \\ &\quad + \int_{-\tau}^{\max\{\theta-t,-\tau\}} g(\chi^0(\gamma-\theta+t),u(\gamma-\theta+t),\gamma)d\mu(\gamma) \\ &\quad + \int_{-\tau}^{\max\{\theta-t,-\tau\}} p(\gamma)d_\gamma y^0(\gamma-\theta+t).\end{aligned} \tag{3.10}$$

At $\theta = 0$, we have

$$\begin{aligned}&\int_{-\tau}^{\max\{-\tau,-t\}} b(\chi^0(\gamma+t),u(t+\gamma),\gamma)d\gamma. \\ &\quad + \int_{-\tau}^{\max\{-t,-\tau\}} g(\chi^0(\gamma+t),u(\gamma+t),\gamma)d\mu(\gamma) \\ &\quad + \int_{-\tau}^{\max\{-t,-\tau\}} p(\gamma)d_\gamma y^0(\gamma+t).\end{aligned} \tag{3.11}$$

Continuing with the main term in (3.6) involving $b(\cdot)$, we have, with the substitution $\gamma = \theta - t + s$ or $s = \gamma - \theta + t$,

$$\begin{aligned}
\int_0^t \Phi(t-s) b(\chi^0(s), u(s), \theta) ds &= \int_0^t b(\chi^0(s), u(s), \theta - t + s) I_{\{-\tau \le \theta - t + s \le 0\}} ds \\
&= \int_{\theta - t}^{\theta} b(\chi^0(\gamma - \theta + t), u(\gamma - \theta + t), \gamma) I_{\{-\tau \le \gamma \le 0\}} d\gamma \\
&= \int_{\max\{\theta - t, -\tau\}}^{\theta} b(\chi^0(\gamma - \theta + t), u(\gamma - \theta + t), \gamma) d\gamma.
\end{aligned} \tag{3.12}$$

Now consider the delayed reflection term. For the main term in (3.6),

$$\begin{aligned}
&\int_0^t \Phi(t-s) p(\theta) dy^0(s) = \int_0^t p(\theta - t + s) I_{\{-\tau \le \theta - t + s \le 0\}} dy^0(s) \\
&= \int_{\theta - t}^{\theta} p(\gamma) I_{\{-\tau \le \gamma \le 0\}} d_\gamma y^0(\gamma - \theta + t) = \int_{\max\{\theta - t, -\tau\}}^{\theta} p(\gamma) d_\gamma y^0(\gamma - \theta + t).
\end{aligned} \tag{3.13}$$

The main term in (3.6) involving the measure $\mu(\cdot)$ was evaluated in (3.7). Adding (3.10), (3.12), (3.13) and (3.7), yields (3.9). Setting $\theta = 0$ in (3.9) and substituting it into (3.5) yields

$$\begin{aligned}
&d\chi^0(t) = dt \int_{-\tau}^0 b(x(t+\gamma), u(t+\gamma), \gamma) d\gamma + dt \int_{-\tau}^0 p(\gamma) d_\gamma y^0(t+\gamma) \\
&+ dt \int_{-\tau}^0 g(\chi^0(t+\gamma), u(t+\gamma), \gamma) d\mu(\gamma) + c(\chi^0(t), u(t)) dt + \sigma(\chi^0(t)) dw(t) + dz^0(t),
\end{aligned}$$

which is the equation for $x(\cdot)$. ■

4 A Discrete Time and State Approximation: Motivation

A numerical procedure. The forms (3.1), (3.2) will motivate the numerical algorithms. But we will need to show that the processes associated with the numerical algorithms converge to (3.5), (3.6) with $\chi^0(\cdot) = x(\cdot)$, where $x(\cdot)$ solves (2.2) or (2.3). The actual numerical algorithms for getting the optimal costs will be discussed in Section 6. To prepare ourselves for that discussion and the types of algebraic manipulations that will be needed, in this section we will discuss a simple discrete time approximation to (3.5), (3.6). This approximation is not intended to be used to solve the optimal control problem, but it will provide helpful insights and guides and is useful for simulation.

Let us formally consider a simple discrete-time discrete-θ approximation to (3.1), (3.2), where δ is the interval for θ and Δ is the time interval and use piecewise constant interpolations in time. Then, for $-\tau < \theta \le 0$, $t = n\Delta$, and letting $\chi^{0,\delta,\Delta}(\cdot), \chi^{1,\delta,\Delta}(\cdot)$ denote the approximations, we can write

$$\begin{aligned}
&\chi^{0.\delta,\Delta}(t+\Delta) - \chi^{0.\delta,\Delta}(t) = \Delta\chi^{1,\delta,\Delta}(t,0) + \Delta c(\chi^{0,\delta,\Delta}(t), u(t)) + \\
&\qquad \sigma(\chi^{0,\delta,\Delta}(t))[w(t+\Delta) - w(t)] + [z^{0,\delta,\Delta}(t+\Delta) - z^{0,\delta,\Delta}(t)], \\
&\chi^{1,\delta,\Delta}(t+\Delta,\theta) - \chi^{1,\delta,\Delta}(t,\theta) = -\left[\chi^{1,\delta,\Delta}(t,\theta) - \chi^{1,\delta,\Delta}(t,\theta-\delta)\right]\frac{\Delta}{\delta} \\
&\qquad\quad +\Delta\left[b(\chi^{0,\delta,\Delta}(t), u(t),\theta) + p(\theta)\left[y^{0,\delta,\Delta}(t+\Delta) - y^{0,\delta,\Delta}(t)\right]\right] \\
&\qquad\quad +g(\chi^{0,\delta,\Delta}(t), u(t),\theta)\left[\mu(\theta) - \mu(\theta-\delta)\right],
\end{aligned} \tag{4.1}$$

with boundary condition $\chi^{1,\delta,\Delta}(t,-\tau) = 0$ and initial condition being the discretization of (3.3). The $z^{0,\delta,\Delta}(\cdot)$ above is the reflection process for $\chi^{0,\delta,\Delta}(\cdot)$, and $d_i y_i^{0,\delta,\Delta}(\cdot)$ is the component due to reflection on the ith face of G. If this process ever leaves the set G, then it is immediately reflected back in accordance with the local reflection direction as defined in (A2.1), (A2.2). Note that the backward difference in θ is used. Since we are interested in only the general approach, let us suppose that $u(\cdot)$ is continuous. In Sections 5 and 6, the controls are arbitrary.

We must have $\Delta = \delta$ if there is to be convergence to the correct limit as the discretization levels go to zero. So, let $\Delta = \delta$ henceforth and with $u^\delta(\cdot)$ denoting the piecewise constant discretization of $u(\cdot)$, rewrite (4.1) as

$$\begin{aligned}
\chi^{0,\delta}(t+\delta) = \chi^{0,\delta}(t) + \delta\chi^{1,\delta}(t,0) + \delta c(\chi^{0,\delta}(t), u^\delta(t)) \\
+\sigma(\chi^{0,\delta}(t))\left[w(t+\delta) - w(t)\right] + \left[z^{0,\delta}(t+\delta) - z^{0,\delta}(t)\right],
\end{aligned} \tag{4.2}$$

$$\begin{aligned}
&\chi^{1,\delta}(t+\delta,\theta) = \chi^{1,\delta}(t,\theta-\delta) + \delta\left[b(\chi^{0,\delta}(t), u^\delta(t),\theta)\right] \\
&\qquad +g(\chi^0(t), u^\delta(t),\theta)\left[\mu(\theta) - \mu(\theta-\delta)\right] + p(\theta)\left[y^{0,\delta}(t+\delta) - y^{0,\delta}(t)\right],
\end{aligned} \tag{4.3}$$

with $\chi^{1,\delta}(t,-\tau) = 0$.

For functions $f(t,\theta)$, define the operator Φ^δ, analogously to (3.4), by

$$(\Phi^\delta f(\cdot))(t,\theta) = f(t,\theta-\delta), \quad -\tau \le \theta - \delta \le 0, \tag{4.4}$$

with $(\Phi^\delta f(\cdot))(t,\theta) = 0$ otherwise. Iterating (4.4) yields

$$[\Phi^\delta]^k f(i\delta,\theta) = f(i\delta,\theta - k\delta) I_{\{-\tau\le\theta-k\delta\le 0\}}. \tag{4.5}$$

Now (4.3) can be written as

$$\begin{aligned}
&\chi^{1,\delta}(t+\delta,\cdot) = \Phi^\delta\chi^{1,\delta}(t,\cdot) + \delta\left[b(\chi^{0,\delta}(t), u^\delta(t),\cdot)\right] \\
&\qquad +g(\chi^0(t), u^\delta(t),\theta)\left[\mu(\theta) - \mu(\theta-\delta)\right] + p(\cdot)\left[y^{0,\delta}(t+\delta) - y^{0,\delta}(t)\right],
\end{aligned} \tag{4.6}$$

with $\chi^{1,\delta}(t,-\tau)=0$. Iterating yields

$$\begin{aligned}\chi^{1,\delta}(n\delta,\theta) = [\Phi^\delta]^n \chi^{1,\delta}(0,\theta) + \sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} \delta \left[b(\chi^{0,\delta}(i\delta), u^\delta(i\delta), \theta) \right] \\ + \sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} \left[g(\chi^{0,\delta}(i\delta), u^\delta(i\delta), \theta) \right] \left[\mu(\theta) - \mu(\theta-\delta) \right] \\ + \sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} p(\theta) \left[y^{0,\delta}(i\delta+\delta) - y^{0,\delta}(i\delta) \right].\end{aligned} \tag{4.7}$$

and

$$\begin{aligned}\chi^{0,\delta}(n\delta) = \chi^{0,\delta}(0) + \delta \sum_{i=0}^{n-1} \chi^{1,\delta}(i\delta, 0) + \delta \sum_{i=0}^{n-1} c(\chi^{0,\delta}(i\delta), u^\delta(i\delta)) \\ + \sum_{i=0}^{n-1} \sigma(\chi^{0,\delta}(i\delta)) \left[w(i\delta+\delta) - w(i\delta) \right] + z^{0,\delta}(n\delta).\end{aligned}$$

Theorem 4.1. *Assume* (A2.1)–(A2.5) *and let the admissible controls* $u^\delta(\cdot)$ *converge to the continuous admissible control* $u(\cdot)$. *Then* $\chi^{0,\delta}(\cdot)$ *converges to* $x(\cdot)$ *and the costs converge as well.*

Proof. Let $\chi^{0,\delta}(\cdot), u^\delta(\cdot)$ denote piecewise constant, right continuous, interpolations, with intervals δ. Let $\epsilon(t,\delta)$ denote a function that goes to zero, uniformly in (t,ω) on any bounded t-interval, as $\delta \to 0$. Its value might change from usage to usage. Then, for $t = n\delta$, the first sum in (4.7) is

$$\begin{aligned}&\delta \sum_{i=0}^{n-1} b(\chi^{0,\delta}(i\delta), u^\delta(i\delta), \theta - t + i\delta + \delta) I_{\{-\tau \le \theta - t + i\delta + \delta \le 0\}} \\ &\quad = \int_0^t b(\chi^{0,\delta}(s), u^\delta(s), \theta - t + s) I_{\{-\tau \le \theta - t + s \le 0\}} ds + \epsilon(t,\delta) \\ &\quad = \int_{\max\{\theta - t, -\tau\}}^{\theta} b(\chi^{0,\delta}(\gamma + t - \theta), u^\delta(\gamma + t - \theta), \gamma) d\gamma + \epsilon(t,\delta).\end{aligned} \tag{4.8}$$

Since the $\chi^{0,\delta}(\cdot)$ and $u^\delta(\cdot)$ are piecewise constant on the intervals $[k\delta, k\delta+\delta)$ and θ is a (negative) integral multiple of δ, the $\epsilon(t,\delta)$ term arises from the θ dependence of $b(x,u,\theta)$. The treatment of the second sum in (4.7) is similar, as follows. For $t = n\delta$,

$$\begin{aligned}&\sum_{i=0}^{n-1} g(\chi^{0,\delta}(i\delta), u^\delta(i\delta), \theta - t + i\delta + \delta) I_{\{-\tau \le \theta - t + i\delta + \delta \le 0\}} \left[\mu(\theta - t + i\delta + \delta) - \mu(\theta - t + i\delta) \right] \\ &= \int_0^t g(\chi^{0,\delta}(s), u^\delta(s), \theta - t + s) I_{\{-\tau \le \theta - t + s \le 0\}} \left[\mu(\theta - t + s) - \mu(\theta - t + s - ds) \right] + \epsilon(t,\delta) \\ &= \int_{\max\{\theta - t, -\tau\}}^{\theta} g(\chi^{0,\delta}(\gamma + t - \theta), u^\delta(\gamma + t - \theta), \gamma) \mu(d\gamma) + \epsilon(t,\delta).\end{aligned} \tag{4.9}$$

The contribution of the last sum in (4.7) is

$$\begin{aligned}\sum_{i=0}^{n-1} p(\theta-t+i\delta+\delta)\left[y^{0,\delta}(i\delta+\delta)-y^{0,\delta}(i\delta)\right] I_{\{-\tau\le\theta-t+i\delta+\delta\le 0\}} \\ =\int_{\max\{\theta-t,-\tau\}}^{\theta} p(\gamma)d_\gamma y^{0,\delta}(\gamma+t-\theta),\end{aligned} \tag{4.10}$$

modulo an error that is due to the approximation of $p(\cdot)$ by a piecewise constant function and that is bounded by $\epsilon(t,\delta)$ times the variation of $y^{0,\delta}(\cdot)$ on the interval $[t-\tau, t-\tau_0]$, where $\tau > \tau_0 > 0$ is defined in (A2.3). The initial condition is treated similarly and, analogously to the development in Theorem 3.1, adds terms of the form $\int_{-\tau}^{\max\{\theta-t,-\tau\}}$ to the expressions computed above.

Putting the pieces together yields

$$\begin{aligned}\chi^{1,\delta}(t,\theta) &= \int_{-\tau}^{\theta} b(\chi^{0,\delta}(t+\gamma-\theta), u^\delta(t+\gamma-\theta),\gamma)d\gamma + \int_{-\tau}^{\theta} p(\gamma)d_\gamma y^{0,\delta}(t+\gamma-\theta) \\ &\quad + \int_{-\tau}^{\theta} g(\chi^{0,\delta}(t+\gamma-\theta), u^\delta(t+\gamma-\theta),\gamma)d\mu(\gamma) \\ &\quad + \epsilon(t,\delta)\left[1+|z^{0,\delta}|(t-\tau_0)-|z^{0,\delta}|(t-\tau)\right],\end{aligned}$$

$$\begin{aligned}\chi^{1,\delta}(t,0) &= \int_{-\tau}^{0} b(\chi^{0,\delta}(t+\gamma), u^\delta(t+\gamma),\gamma)d\gamma + \int_{-\tau}^{0} p(\gamma)d_\gamma y^{0,\delta}(t+\gamma) \\ &\quad + \int_{-\tau}^{0} g(\chi^{0,\delta}(t+\gamma), u^\delta(t+\gamma),\gamma)d\mu(\gamma) + \epsilon(t,\delta)\left[1+|z^{0,\delta}|(t-\tau_0)-|z^{0,\delta}|(t-\tau)\right],\end{aligned} \tag{4.11}$$

$$\begin{aligned}\chi^{0,\delta}(t) &= x(0) + \int_0^t \chi^{1,\delta}(s,0)ds \\ &\quad + \int_0^t c(\chi^{0,\delta}(s), u^\delta(s))ds + \int_0^t \sigma(\chi^{0,\delta}(s))dw(t) + z^{0,\delta}(t).\end{aligned} \tag{4.12}$$

Substituting (4.11) into (4.12) yields that $\chi^{0,\delta}(\cdot)$ satisfies (2.2), modulo the error terms. It follows from Lemma 2.1 and its extension that the error terms go to zero as $\delta \to 0$ in that for each $T < \infty$

$$\lim_{\delta\to 0} E \sup_{t\le T} \epsilon(t,\delta)\left[1+|z^{0,\delta}|(t-\tau_0)-|z^{0,\delta}|(t-\tau)\right] = 0.$$

The main issue concerns the convergence properties of the reflection process $z^{0,\delta}(\cdot)$. A weak convergence argument can be used. The set of processes $(\chi^{0,\delta}(\cdot), u^\delta(\cdot), w(\cdot), z^{0,\delta}(\cdot))$ is tight in the Skorohod topology and all limits are continuous. One needs to show that the weak sense limit $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{w}(\cdot), \bar{z}(\cdot))$ of any weakly convergent subsequence satisfies (2.2), where $\bar{z}(\cdot)$ is the reflection term, and all processes are nonanticipative with respect to the standard Wiener process $\bar{w}(\cdot)$. The proof of [10, Theorem 1.2, Chapter 11] can be used to complete the demonstration. It yields that $\bar{z}(\cdot)$ must be the reflection process for $\bar{x}(\cdot)$, as well as the nonanticipativity properties. Clearly, $(\bar{u}(\cdot), \bar{w}(\cdot))$ has the same probability law as $(u(\cdot), w(\cdot))$. This and the weak-sense uniqueness of solutions implies that the original sequence converges, as desired.

The costs converge due to this pathwise convergence and the continuity of the function $k(\cdot)$, and to the fact that, for any $T < \infty$, Lemma 2.1 and its extension imply that the reflection terms on the intervals $[nT, nT + T]$ are uniformly (in n, δ) integrable, since they satisfy (2.4) and (2.5) uniformly in (small) δ. ■

5 The Markov Chain Approximation Method

5.1 Brief Review of the No-Delay Case

In preparation for the approximation for the delay case, let us recall the basic numerical procedure for the non-delayed problem. The first step is the determination of a finite-state controlled Markov chain that has a continuous-time interpolation that is an "approximation" of the controlled reflected diffusion process $x(\cdot)$. The second step solves the optimization problem for the chain and a cost function that approximates the one used for $x(\cdot)$. Let h denote the approximation parameter. The reference [6] contains some additional details concerning the delay problem. The system of concern in this subsection is

$$dx = b(x, u)dt + \sigma(x)dw + dz. \tag{5.1}$$

where $x \in G$, and $z(\cdot)$ is the reflection process.

To construct the approximation, start by defining S_h, a discretization of $I\!R^r$, which we let be a grid with the distance between points in any coordinate direction being h. The precise requirements are quite general and are spelled out via the local consistency condition given below. In general, the interval can depend on the coordinate direction and the discretized state space need not be a grid. It is only the points in G and their immediate neighbors that will be of interest. Next define the approximating controlled Markov chain ξ_n^h. Its state space will be a subset of S_h, and is usually divided into two parts. The first part is $G_h = G \cap S_h$, on which the chain approximates the diffusion part of (5.1). If the chain tries to leave G_h, then it is returned immediately, consistently with the local reflection direction. Thus, define ∂G_h^+ to be the set of points not in G_h to which the chain might move in one step from some point in G_h. The set ∂G_h^+ is the reflecting boundary for the chain. Let $p^h(x, \tilde{x}|\alpha)$ denote the transition probabilities at state x under control value α.

Local consistency on G_h. Let u_n^h denote the controls used at step n for ξ_n^h. Let $E_{x,n}^{h,\alpha}$ (respectively, $\text{covar}_{x,n}^{h,\alpha}$) denote the expectation (respectively, the covariance) given all of the data to step n, when $\xi_n^h = x, u_n^h = \alpha$. Then the chain must satisfy the following condition. There is a function $\Delta t^h(x, \alpha) > 0$

such that (the formulas define $b^h(\cdot)$ and $a^h(\cdot)$)

$$
\begin{gathered}
E_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] = b^h(x,\alpha)\Delta t^h(x,\alpha) = b(x,\alpha)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)), \\
\text{covar}_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] = a^h(x,\alpha)\Delta t^h(x,\alpha) = a(x)\Delta t^h(x,\alpha) + o(\Delta t^h(x,\alpha)) \\
a(x) = \sigma(x)\sigma'(x), \quad \lim_{h\to 0}\sup_{x,\alpha}\Delta t^h(x,\alpha) = 0, \\
\|\xi_{n+1}^h - \xi_h^n\| \le K_1 h,
\end{gathered}
\tag{5.2}
$$

for some real K_1. With the methods in [10], $\Delta t^h(\cdot)$ is obtained automatically as a byproduct of getting the transition probabilities. Thus, in G, the first two "conditional" moments of $\xi_{n+1}^h - \xi_n^h$ are very close to those of the "differences" of the $x(\cdot)$ of (5.1).[3] Approximations satisfying the local consistency condition (5.2) are to be called *explicit* approximations.

Local consistency on the reflecting boundary ∂G_h^+. From points in ∂G_h^+, the transitions of the chain are such that they move to G_h, with the conditional mean direction being a reflection direction at x. More precisely, $\lim_{h\to 0}\text{distance}(\partial G_h^+, G_h) = 0$, and there are $\theta_1 > 0$ and $\theta_2(h) \to 0$ as $h \to 0$ such that for all $x \in \partial G_h^+$,

$$
\begin{gathered}
E_{x,n}^{h,\alpha}\left[\xi_{n+1}^h - x\right] \in \left\{a\gamma : \gamma \in d(x) \text{ and } \theta_2(h) \ge a \ge \theta_1 h\right\}, \\
\Delta t^h(x,\alpha) = 0 \;\text{ for } x \in \partial G_h^+.
\end{gathered}
\tag{5.3}
$$

The last line of (5.3) says that the reflection from states on ∂G_h^+ is instantaneous. The reference [10] has an extensive discussion of straightforward methods of getting good approximations.

In all cases in [10], the transition probability for the chain in the no-delay case can be represented as a ratio of the following form. For $x \in G_h$,

$$
\begin{gathered}
P\{\xi_1^h = \tilde{x}|\xi_0^h = x, u_0^h = \alpha\} = p^h(x,\tilde{x}|\alpha) = N^h(x,\tilde{x},\alpha)/D^h(x,\alpha), \\
\Delta t^h(x,\alpha) = h^2/D^h(x,\alpha), \quad \inf_{x,\alpha} D^h(x,\alpha) > 0, \quad |\tilde{x} - x| = O(h)
\end{gathered}
\tag{5.4}
$$

where $N^h(\cdot), D^h(\cdot)$ are functions of $b(\cdot), \sigma(\cdot)$ of the form :

$$
N^h(x,\tilde{x},\alpha) = \bar{N}(hb(x,\alpha),\sigma(x),\tilde{x}), \quad D^h(x,\alpha) = \bar{D}(hb(x,\alpha),\sigma(x)). \tag{5.5}
$$

The same general canonical forms (5.4) and (5.5) will be used for the delay case. Define $\Delta t_n^h = \Delta t^h(\xi_n^h, u_n^h)$, $\delta z_n^h = [\xi_{n+1}^h - \xi_n^h] I_{\{\xi_n^h \notin G\}}$ and let $\mathcal{F}_n^h$ denote the minimal σ-algebra that measures the system data to step n. By centering around the conditional expectation, we can write

$$
\xi_{n+1}^h = \xi_n^h + \Delta t_n^h b(\xi_n^h, u_n^h) + \beta_n^h + \delta z_n^h + o(\Delta t_n^h), \tag{5.6}
$$

[3] The consistency condition need not hold everywhere. See [10, discussion in Section 5.5 and Theorem 10.5.3, and also the discussion concerning discontinuous dynamics in Section 10.2 of second ed.] for examples and more detail.

where the martingale difference β_n^h has conditional (on $\mathcal{F}_n^h$) covariance $a^h(\xi_n^h, u_n^h)\Delta t_n^h$. The $o(\Delta t_n^h)$ term is due to the use of $b(\cdot)$ in lieu of $b^h(\cdot)$ since by (5.2),

$$\lim_{h\to 0} \sup_{x,\alpha} |b^h(x,\alpha) - b(x,\alpha)| = 0.$$

.

Continuous-time interpolations. The discrete-time chain ξ_n^h is used for the numerical computations. However, for the proofs of convergence, we use a continuous-time interpolation $\psi^h(\cdot)$ that will approximate $x(\cdot)$. This is constructed as follows. Let $\nu_n, n = 0, 1, \dots,$ be mutually independent and exponentially distributed random variables with unit mean, and that are independent of $\{\xi_n^h, u_n^h, n \geq 0\}$. Define $t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h$, $\Delta\tau_n^h = \nu_n \Delta t_n^h$, and $\tau_n^h = \sum_{i=0}^{n-1} \Delta\tau_i^h$. Then define $\psi^h(t) = \xi_n^h$ for $t \in [\tau_n^h, \tau_{n+1}^h)$. Define the continuous-time interpolations $u^h(\cdot)$ of the control actions analogously and let its relaxed control representation be denoted by $m^h(\cdot)$, with time derivative $m^{h,\prime}(\cdot)$. Let $z^h(\cdot)$ denote the interpolation of $\sum_{i=0}^{n-1} \delta z_i^h$.

Since the intervals between jumps are $\Delta t_n^h \nu_n$, where ν_n is exponentially distributed and independent of $\mathcal{F}_n^h$, the jump rate of $\psi^h(\cdot)$ when in state x is $1/\Delta t^h(x)$. Given a jump, the distribution of the next state is given by the $p^h(x, y|\alpha)$, and the conditional mean change, for $x \in G_h$ and control value α used, is $b^h(x,\alpha)\Delta t^h(x,\alpha)$. So we can decompose $\psi^h(\cdot)$ in terms of the continuous-time compensator, reflection term, and and martingale as

$$\psi^h(t) = x(0) + \int_0^t b^h(\psi^h(s), u^h(s))ds + M^h(t) + z^h(t), \tag{5.7}$$

where the martingale $M^h(t)$ has quadratic variation process $\int_0^t a^h(\xi^h(s), u^h(s))ds$. In terms of the relaxed control we can write

$$\int_0^t b^h(\psi^h(s), u^h(s))ds = \int_0^t \int_U b^h(\psi^h(s), \alpha) m^{h,\prime}(s, d\alpha) ds.$$

It can be shown that ([10, Section 10.4.1]) there is a martingale $w^h(\cdot)$ (with respect to the filtration generated by the state and control processes, possibly augmented by an "independent" Wiener process) such that

$$M^h(t) = \int_0^t \sigma^h(\xi^h(s), u^h(s)) dw^h(s) = \int_0^t \sigma(\xi^h(s)) dw^h(s) + \epsilon^h(t),$$

where $\sigma^h(\cdot)[\sigma^h(\cdot)]' = a^h(\cdot)$ (recall the definition of $a^h(\cdot)$ in (5.2)), $w^h(\cdot)$ has quadratic variation process It and converges weakly to a standard Wiener process. The martingale $\epsilon^h(\cdot)$ is due to the difference between $\sigma(x)$ and $\sigma^h(x,\alpha)$ and

$$\lim_{h\to 0} \sup_{u^h} \sup_{s\leq t} E|\epsilon^h(s)|^2 = 0 \tag{5.8}$$

for each t. Thus

$$\psi^h(t) = x(0) + \int_0^t \int_U b^h(\psi^h(s), \alpha) m^{h,\prime}(d\alpha, s) ds + \int_0^t \sigma(\psi^h(s)) dw^h(s) + z^h(t) + \epsilon^h(t). \tag{5.9}$$

The following result [10, Theorem 1.3, Chapter 11] is an analog of Lemma 2.1.

Lemma 5.1. *Assume* (A2.1)–(A2.2) *and that for some constant K, the $b(x)$ and $\sigma(x)$ in* (5.2) *are replaced by arbitrary $\mathcal{F}_n^h$-measurable random variables that are bounded in norm by K. Then the corresponding reflection terms satisfy*

$$\lim_{T \to 0} \limsup_{h \to 0} \sup_{b(\cdot), \sigma(\cdot), x(0)} E \left| z^h \right|^2 (T) = 0.$$

For any $T < \infty$,

$$\limsup_{h \to 0} \sup_{b(\cdot), \sigma(\cdot), x(0)} E \left| z^h \right|^2 (T) < \infty.$$

Note on convergence. For any subsequence $h \to 0$, there is a further subsequence (also indexed by h for simplicity) such that $(\psi^h(\cdot), m^h(\cdot), w^h(\cdot), z^h(\cdot))$ converges weakly to random processes $(x(\cdot), m(\cdot), w(\cdot), z(\cdot))$, where $m(\cdot)$ is a relaxed control, $(x(\cdot), m(\cdot), w(\cdot), z(\cdot))$ is nonanticipative with respect to the standard vector-valued Wiener process $w(\cdot)$, and the set satisfies

$$x(t) = x(0) + \int_0^t \int_U b(x(s), \alpha) m'(d\alpha, s) ds + \int_0^t \sigma(x(s)) dw(s) + z(t),$$

where $z(\cdot)$ is the reflection term. Along the selected subsequence $W^h(x, m^h) \to W(x, m)$. The proofs of these facts are in [10, Chapters 10, 11].

5.2 The "Implicit" Numerical Procedure

It was noted in Section 4 that, for the approximations used there, the discretization levels for time and for the θ variable must be the same. The time interval implied by the Markov chain approximation discussed above is $\Delta t^h(\xi_n^h, u_n^h) = \Delta t_n^h$, which is commonly of the order of h^2, but is not usually a constant. The transition probabilities can be modified so that $\Delta t_n^h = \Delta t^h$, a constant, also of order $O(h^2)$. If we used a discretization of the θ variable with levels $O(h^2)$, there would be far too many θ-values for practical use. An alternative approximation, known as the implicit method [10, Chapter 12], avoids such problems and will be readily adapted to and particularly useful for the delay problem. The fundamental difference between the approximation discussed above and the so-called implicit approaches to the Markov chain approximation lies in the fact that in the former the time variable is treated differently than the state variables: It is a true "time" variable, and its value increases by Δt_n^h at step n. In the implicit approach, the time variable is treated as just another state variable. It is discretized in the same manner as are the other state variables: For the no-delay case, the approximating Markov chain has a state space that is a discretization of the $(x,t)-$space, and the component of the state of the chain that comes from the original time variable does not necessarily increase its value at each step. The idea is analogous when there are delays. Let $\delta > 0$ be the discretization level for the time variable. We will assume that $\delta = O(h)$. Let $p^{h,\delta}(x, n\delta; \tilde{x}, n\delta|\alpha)$ denote the probability that the state x moves to $\tilde{x}$, and the time variable does not advance, under control α. Let $p^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)$ denote the probability that the state does not change but that the time advances.

Given the transition probabilities and interpolation interval $p^h(\cdot), \Delta t^h(\cdot)$, those for the implicit method can readily be computed [10, Section 12.4]. Suppose that at the current step the time variable does not advance. Then, conditioned on this event and on the value of the current spatial state, the distribution of the next spatial state is just the $p^h(x, \tilde{x}|\alpha)$ used previously. So one needs only determine the conditional probability that the time variable advances, conditioned on the current state. This is obtained by a "local consistency" argument and no matter how the $p^h(\cdot)$ were derived, the (no-delay) transition probabilities $p^{h,\delta}(\cdot)$ and interpolation interval $\Delta t^{h,\delta}(\cdot)$ for the implicit procedure can be determined from the $p^h(\cdot), \Delta t^h(\cdot)$ by the formulas [10, Section 12.4], for $x \in G_h$,

$$\begin{aligned} p^h(x, \tilde{x}|\alpha) &= \frac{p^{h,\delta}(x, n\delta; \tilde{x}, n\delta|\alpha)}{1 - p^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha)}, \\ p^{h,\delta}(x, n\delta; x, n\delta + \delta|\alpha) &= \frac{\Delta t^h(x, \alpha)}{\Delta t^h(x, \alpha) + \delta}, \\ \Delta t^{h,\delta}(x, \alpha) &= \frac{\delta \Delta t^h(x, \alpha)}{\Delta t^h(x, \alpha) + \delta}. \end{aligned} \tag{5.10}$$

The reflecting states are treated as before. One usually first computes the transition functions for the explicit case, where they must satisfy (5.2), and then uses those to get the transition probabilities for the implicit case via (5.10).

Let $\xi_n^{h,\delta}$ and $\phi_n^{h,\delta}$ denote the spatial and temporal states at step n under (5.10). Define $\Delta t_n^{h,\delta} = \Delta t^{h,\delta}(\xi_n^{h,\delta}, u_n^{h,\delta})$. Under (5.10), (5.2) holds for the process $\xi_n^{h,\delta}$, with the time interval $\Delta t^{h,\delta}(\cdot)$. We have $E_n^h[\phi_{n+1}^{h,\delta} - \phi_n^{h,\delta}] = \Delta t_n^{h,\delta}$ and

$$\phi_{n+1}^{h,\delta} = \phi_n^{h,\delta} + \Delta t_n^{h,\delta} + \beta_{0,n}^{h,\delta}. \tag{5.11}$$

where the covariance of the martingale difference term $\beta_{0,n}^{h,\delta}$ is $o(\Delta t_n^{h,\delta})$. Let $\delta z_n^{h,\delta}$ denote the reflection term for the $\xi_n^{h,\delta}$ process, and define the components $\delta y_{i,n}^{h,\delta}$ by $\delta z_n^{h,\delta} = \sum_i d_i \delta y_{i,n}^{h,\delta}$. Analogously to what was done for the explicit case, define $\Delta \tau_n^{h,\delta} = \nu_n \Delta t_n^{h,\Delta}$, and $\tau_n^{h,\Delta} = \sum_{i=0}^{n-1} \Delta \tau_i^{h,\Delta}$. The continuous time interpolation is $\psi^{h,\delta}(t) = \xi_n^{h,\delta}$ for $t \in [\tau_n^{h,\delta}, \tau_{n+1}^{h,\delta})$, with the analogous interpolations $u^{h,\delta}(\cdot)$ and $z^{h,\delta}(\cdot)$ for the controls and reflection term $\sum_{i=0}^{n-1} \delta z_i^{h,\delta}$, resp.

Asymptotic equivalence of the time scales. For each $t \geq 0$, define the stopping times

$$d^h(t) = \max\left\{n : \sum_{i=0}^{n-1} \Delta t_i^h = t_n^h \le t\right\}, \quad d^{h,\delta}(t) = \max\left\{n : \sum_{i=0}^{n-1} \Delta t_i^{h,\delta} = t_n^{h,\delta} \le t\right\}.$$

Note that $d^h(t)$ will *never* be the index of a reflecting state, since the time intervals for those are zero. Define $d_\tau^h(t)$ and $d_\tau^{h,\delta}(t)$ analogously, but with $\Delta\tau_n^h$ and $\Delta\tau_n^{h,\delta}$ used, resp.

Theorem 5.2. *For each $t > 0$,*

$$\lim_{h\to 0} \sup_{u^h} E \sup_{s\le t} \left[\sum_{i=0}^{d^h(s)} (\Delta\tau_i^h - \Delta t_i^h)\right]^2 = 0. \tag{5.12}$$

Also,

$$\lim_{h\to 0,\delta\to 0} \sup_{u^{h,\delta}} E \sup_{s\le t} \left[\sum_{i=0}^{d^{h,\delta}(s)} (\Delta\tau_i^{h,\delta} - \Delta t_i^{h,\delta})\right]^2 = 0. \tag{5.13}$$

The last assertion holds with $d_\tau^{h,\delta}(\cdot)$ replacing $d^{h,\delta}(\cdot)$ and also with $\Delta t_n^h(1-I_n^{h,\delta})$ replacing $\Delta t_n^{h,\delta}$. Let $\phi^{h,\delta}(\cdot)$ denote the interpolation of the $\phi_n^{h,\delta}$ with the intervals $\Delta t_n^{h,\delta}$. Then $\phi^{h,\delta}(\cdot)$ converges weakly to the process with value t at time t. The result of the last sentence holds if the intervals are $\Delta\tau_n^{h,\delta}$.

Proof. Owing to the mutual independence of the exponential random variables $\{\nu_n\}$ and their independence of everything else, the discrete parameter process $L_n = \sum_{i=0}^n (\Delta\tau_i^h - \Delta t_i^h)$ is a martingale. Hence, the conditional expectation of the squared term in (5.12) given the $\{\Delta t_i^h\}$ equals

$$E\left[\sum_{i=0}^{d^h(t)} \left[\Delta\tau_i^h - \Delta t_i^h\right]^2 \middle| \Delta t_i^h, i < \infty\right] = \sum_{i=0}^{d^h(t)} [\Delta t_i^h]^2 \le (t + \sup_n \Delta t_n^h) \sup_n \Delta t_n^h \xrightarrow{h} 0,$$

which yields (5.12) since by Doob's inequality and the martingale property, $E\sup_{j\le n} |L_j|^2 \le 4E|L_n|^2$. Equation (5.13) and the assertion following it are proved in the same way.

For the next to last assertion of the theorem, write

$$\phi^{h,\delta}(t) = \sum_{i=0}^{d^{h,\delta}(t)-1} \Delta t_i^{h,\delta} + \sum_{i=0}^{d^{h,\delta}(t)-1} \beta_{0,i}^{h,\delta}.$$

The first sum equals t, modulo $\sup_n \Delta t_n^{h,\delta}$. The variance of the martingale term is δt, modulo $\delta + \sup_n \Delta t_n^{h,\delta}$, and the term converges weakly to the zero process. This yields the next to last assertion of the theorem. This and (5.13) yields the last assertion of the theorem. ■

6 Approximating the System With Delays

The numerical approximations and convergence proofs for the delay problem will now be developed. Although the form of the algorithm is motivated by those used for the no-delay problem, it is more complicated. But the delay problem is much more complicated and we are interested in algorithms that ameliorate the memory requirements. It is seen from the discussion in Section 4 that the time variable and θ need to have the same increments if there is to be convergence to the correct values. Discretizing θ and then $\chi^0(t), \chi^1(t,\theta)$ so that (5.2) holds for the resulting process would require time and θ intervals of order $O(h^2)$, so that θ would take $O(1/h^2)$ values, much too high. The implicit method for the Markov chain approximation can be adapted to alleviate these problems.

Let τ be an integral multiple of δ and let θ take values in $T^\delta = \{-\tau + \delta, \ldots, -\delta, 0\}$. Owing to the boundary condition $\chi^1(t,-\tau) = 0$ there is no need for $\theta = -\tau$. The Markov chain approximating $(\chi^0(t), \chi^1(t,\theta), -\tau \le \theta \le 0)$ will be denoted by $(\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(\theta), \theta \in T^\delta)$. The $\xi_n^{0,h,\delta}$ will take values in $G_h \cup \partial G_h^+$ with instantaneous reflection back if it leaves G_h, in accordance with (A2.1) and (A2.2). We suppose that for each $\theta \in T^\delta$, $\xi_n^{1,h,\delta}(\theta)$ takes values in a regular h-grid. Any interval of order $O(h)$ can be used, but the notation becomes more awkward.

Boundaries for $\boldsymbol{\xi_n^{1,h,\delta}(\theta)}$. The $\xi_n^{0,h,\delta}$ are in G, hence bounded. If $p(\cdot) = 0$, then (3.9) shows that $|\chi^1(t,\theta)|$ is bounded by $|b|(\tau+\theta)+|g|\mu([-\tau,\theta])$ where $|b|, |g|$ are the sup values, and $|\xi_n^{1,h,\delta}(\theta)|$ can be taken to be bounded by slightly higher values. One could stop the process $\xi_n^{1,h,\delta}(\theta)$ when it reached these values. In the limit, as $h, \delta \to 0$, the $\xi_n^{1,h,\delta}(\theta)$ would not reach the boundaries and no other precautions need to be taken in the construction of the algorithm. However, if there is a delayed reflection term, then the bound is changed by the middle term in (3.9). This term is bounded by a constant times the variation of $z(\cdot)$ on $[t-\tau_0, t-\tau]$, which satisfies Lemma 5.1 and, in the limit, Lemma 2.1. For numerical purposes, we need to bound $|\xi_n^{1,h,\delta}(\theta)|$. This will be done by stopping it when it reaches some level such that the probability that the level is exceeded by $|\chi^1(t,\theta)|$ on some large time interval $[0,T]$ is small. The $\xi_n^{1,h,\delta}(\theta)$ will be stopped for all θ if any one hits the boundary. The required bound depends on the discount factor β and the value of τ. If the bound is reached for small h, δ, then there will be an error in computing the cost, the error depending on the value of $e^{-\beta t}$ at that time. A reasonable procedure is to have the boundaries at least twice what is needed if there were no delayed reflection term. If a boundary is ever reached, the drop the reflection term and continue with the others. Until Theorem 6.3, ignore these boundaries.

The algorithm for the implicit procedure. As in Section 5 for the so-called implicit method, let $\phi_n^{h,\delta}$ denote the time variable. The algorithm is slightly different from the implicit method defined in Section 5. As there, the steps can be divided into two classes, corresponding to the time variable advancing or not.

Owing to the coordination between the advance in the time variable and the "shift" in θ associated with the operator Φ^δ, if time advances at step n, then the transitions are

$$\xi_{n+1}^{1,h,\delta}(\theta) = \xi_n^{1,h,\delta}(\theta - \delta), \quad \xi_{n+1}^{0,h,\delta} = \xi_n^{0,h,\delta}, \quad \phi_{n+1}^{h,\delta} = \phi_n^{h,\delta} + \delta. \tag{6.1}$$

The transition in $(\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(\theta), \theta \in T^\delta)$ if the time variable does not advance at step n will be defined below. Let $u_n^{h,\delta}$ denote the control applied at step n and let $\mathcal{F}_n^{h,\delta}$ denote the minimal σ-algebra that measures the system data to step n, with the associated conditional expectation denoted by $E_n^{h,\delta}$.

Recall the canonical form of the transition probabilities and time interval in (5.4), (5.5). The identical forms will be used here for the transitions of the component $\xi_n^{0,h,\delta}$. Thus, by (5.4), (5.5), for $\xi_n^{0,h,\delta} = x^0 \in G_h, \xi_n^{1,h,\delta}(0) = x^1$ and control value α used, the probability that $\xi_{n+1}^{0,h,\delta}$ takes the value $\tilde{x}^0$, conditioned on the event that the time variable does not advance at step n, is

$$p^h\left(x^0, \tilde{x}^0 \middle| \alpha, x^1\right) = \frac{\bar{N}\left(h(x^1 + c(x^0,\alpha)), \sigma(x^0), \tilde{x}^0\right)}{\bar{D}\left(h(x^1 + c(x^0,\alpha)), \sigma(x^0)\right)}, \quad \left|\tilde{x}^0 - x^0\right| = O(h). \tag{6.2}$$

Thus the transition probability for $\xi_n^{0,h,\delta}$ has the same dependence on the full drift vector and covariance matrix as for the non-delay case. Indeed, with the correct drift and covariance used, any of the algorithms in [10, Chapter 5] can be used to get (6.2).

Next define the interval $\Delta t^h(x^0, x^1, \alpha) = h^2/[\bar{D}(h(x^1 + c(x^0,\alpha)), \sigma(x^0))]$. Define the probability that the time variable advances at step n by (5.10), namely,

$$p^{h,\delta}(x^0, n\delta; x^0, n\delta + \delta | \alpha, x^1) = \frac{\Delta t^h(x^0, x^1, \alpha)}{\Delta t^h(x^0, x^1, \alpha) + \delta}. \tag{6.3}$$

Also, from (5.10), define the interval for the implicit procedure:

$$\Delta t^{h,\delta}(x^0, x^1, \alpha) = \frac{\delta \Delta t^h(x^0, x^1, \alpha)}{\Delta t^h(x^0, x^1, \alpha) + \delta}.$$

If $\xi_n^{0,h,\delta} = x^0 \notin G_h$, then the reflection back to G_h is instantaneous in that $\Delta t^h(x^0, x^1, \alpha) = 0$, and it is in accord with (A2.1)-(A2.2). Set $\delta z_n^{0,h,\delta} = [\xi_{n+1}^{0,h,\delta} - \xi_n^{0,h,\delta}] I_{\{\xi_n^{0,h,\delta} \notin G_h\}}$,

$$\Delta t_n^h = \Delta t^h(\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(0), u_n^{h,\delta}), \quad \Delta t_n^{h,\delta} = \Delta t^{h,\delta}(\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(0), u_n^{h,\delta}).$$

At step n we first decide, according to (6.3), whether time advances or not. The evolution of the time variable can be decomposed as (5.11). Let $I_n^{h,\delta}$ denote the indicator function of the event that the time variable advances at step n. By the local consistency condition, for $\xi_n^{0,h,\delta} \in G_h$ and on the event that time does not advance, we must have

$$\begin{aligned} &E\left[\xi_{n+1}^{0,h,\delta} - \xi_n^{0,h,\delta} \middle| I_n^{h,\delta} = 0, \mathcal{F}_n^{h,\delta}\right] = \Delta t_n^h\left[\xi_n^{1,h,\delta}(0) + c(\xi_n^{0,h,\delta}, u_n^{h,\delta})\right] + o(\Delta t_n^h), \\ &\text{covar}_n^{h,\delta,\alpha}\left[\xi_{n+1}^{0,h,\delta} - \xi_n^{0,h,\delta}\right] = a(\xi_n^{0,h,\delta})\Delta t_n^h + o(\Delta t_n^h). \end{aligned} \tag{6.4}$$

Thus, we can write

$$\xi_{n+1}^{0,h,\delta} = \xi_n^{0,h,\delta} + \left[\Delta t_n^h \xi_n^{1,h,\delta}(0) + \Delta t_n^h c(\xi_n^{0,h,\delta}, u_n^{h,\delta}) + \tilde{\beta}_n^{0,h,\delta} + \delta z_n^{0,h,\delta}\right]\left(1 - I_n^{h,\delta}\right) + o(\Delta t_n^h),$$

where the conditional covariance of the martingale difference $\tilde{\beta}_n^{0,h,\delta}$ is $a(\xi_n^{0,h,\delta})\Delta t_n^h + o(\Delta t_n^h)$. By centering $I_n^{h,\delta}$ about its conditional expectation, the above expression can be written as

$$\xi_{n+1}^{0,h,\delta} = \xi_n^{0,h,\delta} + \Delta t_n^{h,\delta} \xi_n^{1,h,\delta}(0) + \Delta t_n^{h,\delta} c(\xi_n^{0,h,\delta}, u_n^{h,\delta}) + \beta_n^{0,h,\delta} + \delta z_n^{0,h,\delta} + o(\Delta t_n^h), \tag{6.5}$$

where the conditional covariance of the martingale difference $\beta_n^{0,h,\delta}$ is $a(\xi_n^{0,h,\delta})\Delta t_n^{h,\delta} + o(\Delta t_n^{h,\delta})$. Define $\delta\bar{z}_n^{0,h,\delta} = E\left[\delta z_n^{0,h,\delta} \middle| I_n^{h,\delta} = 0, \mathcal{F}_n^{h,\delta}\right]$. Then we can write $\delta\bar{z}_n^{0,h,\delta} = \sum_i d_i \delta\bar{y}_{i,n}^{0,h,\delta}$, where $\delta\bar{y}_{i,n}^{0,h,\delta}$ can only increase if the reflection is to the ith face of G. The error terms $\delta z_n^{0,h,\delta} - \delta\bar{z}_n^{0,h,\delta}$ are asymptotically negligible and will be ignored; see [10, Theorem Equation 11.1.16], where a slightly different notation is used. The values of $\delta\bar{y}_n^{0,h,\delta}$ and $\delta\bar{z}_n^{0,h,\delta}$ are known at step n.

The rule for (6.1) was easy to establish since $\chi^0(t)$ evolves as a diffusion, so any of the algorithms in [10] could be readily adapted. To develop an algorithm for the $\xi_n^{1,h,\delta}(\theta)$ we use the development in Section 4 as a guide. The change between updates of the time variable uses the three right hand terms of (4.3) as a guide. Since there are usually several steps between updates of the time variable, those three terms are approximated as a sum of terms. We proceed as follows.

If time does not advance at step n, then $\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta)$ is to satisfy

$$E\left[\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta) \middle| I_n^{h,\delta} = 0, \mathcal{F}_n^{h,\delta}\right] = q_n^{h,\delta}(\theta), \tag{6.6}$$

where (note that there is no "shift term" in θ since time is not advancing)

$$q_n^{h,\delta}(\theta) = b(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\Delta t_n^h + p(\theta)\delta\bar{y}_n^{0,h,\delta} + g(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\frac{[\mu(\theta) - \mu(\theta - \delta)]}{\delta}\Delta t_n^h,$$

and the boundary condition is $\xi_n^{1,h,\delta}(-\tau) = 0$. To attain the conditional mean (6.6), one randomizes between grid points that are closest to $q_n^{h,\delta}(\theta)$. Write

$$\left[\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta)\right]\left(1 - I_n^{h,\delta}\right) = \left[q_n^{h,\delta}(\theta) + \rho_n^{h,\delta}(\theta)\right]\left(1 - I_n^{h,\delta}\right), \tag{6.7}$$

where $\rho_n^{h,\delta}(\theta)$ is the randomization noise. The randomization can be correlated in θ, but is independent in n. Theorem 6.1 shows that the noise due to the randomization is asymptotically negligible. Recall that if $\delta y_n^{0,h,\delta} \neq 0$, then $\Delta t_n^h = 0$ and conversely.

If we sought to add the terms corresponding to the right hand three terms in (4.3) only at the steps when the time variable is advanced, we would have to keep track of the running sums of the $q_n^{h,\delta}(\theta)$ between such updates, which would amount to an additional state component. No doubt there are other forms that are advantageous.

The cost function is a discretization of (2.6). The initial data $\chi^0(0), \chi^1(0,\cdot)$ for (3.1), (3.2), is the function of the initial path and control segments $\hat{x}, \hat{u}$, resp., as given by (3.3). Without loss of generality, we can suppose that the values of the initial data (3.3) are on the grid for each θ. If control $u(\cdot)$ is used on $[0,\infty)$, then the cost function can be written as

$$
\begin{aligned}
W^{h,\delta}(\chi^0(0), \chi^1(0), u) &= W^{h,\delta}(\hat{x}, \hat{u}, u) \\
&= E^u_{\hat{x},\hat{u}} \sum_{i=0}^{\infty} e^{-\beta t_n^{h,\delta}} \left[k(\xi_n^{0,h,\delta}, u_n^{h,\delta}) \Delta t_n^{h,\delta} + q' \delta \bar{y}_n^{0,h,\delta} \right],
\end{aligned} \tag{6.8}
$$

where $E^u_{\hat{x},\hat{u}}$ denotes the expectation given initial data $\hat{x}, \hat{u}$ and the use of control $u(\cdot)$ on $[0,\infty)$. Let $V^{h,\delta}(\chi^0(0), \chi^1(0)) = V^{h,\delta}(\hat{x}, \hat{u})$ denote the infimum of the costs.

Let $\nu_n, n = 0, 1, \ldots,$ be mutually independent and exponentially distributed random variables with unit mean, and that are independent of $\{\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(\cdot), u_n^h, n \geq 0\}$ and the initial data. As in Section 5. define $\Delta\tau_n^{h,\delta} = \Delta t_n^{h,n} \nu_n$, The expression (6.8) is asymptotically equivalent to

$$
E^u_{\hat{x},\hat{u}} \sum_{i=0}^{\infty} e^{-\beta \tau_n^{h,\delta}} \left[k(\xi_n^{0,h,\delta}, u_n^{h,\delta}) \Delta \tau_n^{h,\delta} + q' \delta \bar{y}_n^{0,h,\delta} \right]. \tag{6.9}
$$

To assure that (6.9) is well defined, note that for any constant $D > 0$ and v being exponentially distributed with mean unity, $Ee^{-vD} = 1/[1 + D]$. This implies that the exponential in (6.9) is equivalent to $\prod_{i=0}^{n-1} 1/[1+\beta\Delta t_i^{h,\delta}]$, which implies that (6.9) is well defined and shows shows the asymptotic equivalence between (6.8) and (6.9) as well. The form (6.8) (with perhaps a numerically convenient approximation to the exponential) would be used in the numerical computations and the form (6.9) in the proofs of convergence.

Recall the continuous time interpolation $\psi^h(\cdot)$ in Section 5. Define $\tau_n^{h,\delta} = \sum_{i=0}^{n-1} \Delta\tau_i^{h,\delta}$. Set $\psi^{0,h,\delta}(t) = \xi_n^{0,h,\delta}$ for $t \in [\tau_n^{h,\delta}, \tau_{n+1}^{h,\delta})$, and define the interpolations $\psi^{1,h,\delta}(t,\theta)$, $u^{h,\delta}(\cdot)$, $z^{0,h,\delta}(\cdot), \bar{z}^{0,h,\delta}(\cdot)$ and $y^{0,h,\delta}(\cdot), \bar{y}^{0,h,\delta}(\cdot)$ analogously. Let $m^{h,\delta}(\cdot)$ be the relaxed control representation of $u^{h,\delta}(\cdot)$, with the derivative at t denoted by $m^{h,\delta,\prime}(\cdot,t)$. Analogously to (5.7) and (5.9) we can write

$$
\begin{aligned}
\psi^{0,h,\delta}(t) = x(0) &+ \int_0^t \xi^{1,h,\delta}(s,0) ds \\
&+ \int_0^t \int_U c(\psi^{0,h,\delta}(s,\alpha) m^{h,\delta,\prime}(d\alpha, s) ds + M^{h,\delta}(t) + z^{0,h,\delta}(t).
\end{aligned}
$$

The process $\bar{z}^{0,h,\delta}(\cdot)$ is asymptotically equivalent to $z^{0,h,\delta}(\cdot)$. There is a martingale $w^{h,\delta}(\cdot)$ with quadratic variation process It and which converges weakly to a Wiener process as $(h,\delta) \to 0$ such that

$$
M^{h,\delta}(t) = \int_0^t \sigma(\psi^{0,h,\delta}(s)) dw^{h,\delta}(s) + \epsilon^{h,\delta}(t),
$$

where $\epsilon^{h,\delta}(\cdot)$, satisfies (5.8).

Define $v_0^{h,\delta} = 0$ and, for $n \geq 0$, set $v_{n+1}^{h,\delta} = \min\{i > v_n^{h,\delta} : \phi_{i+1}^{h,\delta} - \phi_i^{h,\delta} = \delta\}$. The next step is to show that the noise due to the randomization done to attain the conditional mean in (6.6) is negligible. This is (6.10) below. The expression (6.11) shows that $\xi_n^{1,h,\delta}(0)$ changes little *between* the steps that the time variable changes.

Theorem 6.1. *Assume* (A2.1)–(A2.5) *and that* $\Delta t^h(x^0, x^1, \alpha) = O(h^2)$. *Then, for any* $t < \infty$*,*

$$\lim_{h,\delta\to 0} \sup_{u^{h,\delta},\hat{x},\hat{u},\theta} E \sup_{n:\tau_n^{h,\delta}\leq t} \left| \sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} R_i^{h,\delta}(\theta) \right|^2 = 0, \tag{6.10}$$

where[4]

$$R_i^{h,\delta}(\theta) = \sum_{l=v_i^{h,\delta}+1}^{v_{i+1}^{h,\delta}-1} \rho_l^{h,\delta}(\theta),$$

where $\rho_l^{h,\delta}(\theta)$ *is the randomization noise defined in* (6.7)*. Also,*

$$\lim_{h,\delta\to 0} \sup_{u^{h,\delta},\hat{x},\hat{u}} \sup_{n:v_n^{h,\delta}\leq t} E \sup_{v_n^{h,\delta}+1\leq l\leq v_{n+1}^{h,\delta}} \left| \xi_l^{1,h,\delta}(0) - \xi_{v_n^{h,\delta}+1}^{1,h,\delta}(0) \right|^2 = 0. \tag{6.11}$$

Proof. For notational simplicity, write $v_n^{h,\delta}$ simply as v_n. First, let $\xi_n^{0,h,\delta} \notin G_h$, so that we are at a reflection step, and consider the randomization noise associated with realizing the conditional mean $p(\theta)\delta\bar{y}_n^{0,h,\delta}$. Without loss of generality, suppose that it is real-valued. Suppose that $p(\theta)\delta\bar{y}_n^{0,h,\delta}$ lies in $[l_n h, l_n h+h]$ where l_n is a (random) integer, either negative or positive. Then randomize between the end points to get the desired mean $p(\theta)\delta\bar{y}_n^{0,h,\delta}$. To evaluate the variance (always conditioned on $\mathcal{F}_n^{h,\delta}$), without loss of generality we can suppose that we have shifted the means so that $l_n = 0$. Then the probability of selecting h is $p(\theta)\delta\bar{y}_n^{0,h,\delta}/h$ and, since $\delta\bar{y}_n^{0,h,\delta} = O(h)$, the conditional variance is

$$\left[h - p(\theta)\delta\bar{y}_n^{0,h,\delta}\right]^2 \frac{p(\theta)\delta\bar{y}_n^{0,h,\delta}}{h} + [p(\theta)\delta\bar{y}_n^{0,h,\delta}]^2 \left[1 - \frac{p(\theta)\delta\bar{y}_n^{0,h,\delta}}{h}\right] = O(h)|p(\theta)\delta\bar{y}_n^{0,h,\delta}|. \tag{6.12}$$

To evaluate the total effect on the interval $[0, t]$, define the effect on the randomization error due to the reflection steps:

$$Q_n^{h,\delta}(\theta) = \sum_{l=v_n+1}^{v_{n+1}-1} \left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta) \right) - p(\theta)\delta\bar{y}_l^{0,h,\delta} \right] I_{\{\xi_l^{0,h,\delta}\notin G_h\}}.$$

[4]The $\sup_{n:\tau_n^{h,\delta}\leq t}$ is over all stopping times no bigger than the time needed to get to interpolated time t.

Then evaluate

$$\sum_{i=0}^{n-1} [\Phi^{\delta}]^{n-i-1} Q_i^{h,\delta}(\theta), \tag{6.13}$$

which is the total effect of the randomization noise on $\xi_{v_n}^{1,h,\delta}(\theta)$. The summands in the last two expressions are martingale differences since

$$E_{v_n+l}^{h,\delta}\left[\left(\xi_{v_n+l+1}^{1,h,\delta}(\theta) - \xi_{v_n+l}^{1,h,\delta}(\theta)\right) - p(\theta)\delta\bar{y}_{v_n+l}^{0,h,\delta}\right] I_{\{\xi_{v_n+l}^{0,h,\delta}\notin G_h\}} I_{\{v_n+l<v_{n+1}\}} = 0.$$

By this martingale property the mean square value of (6.13) is

$$E\sum_{i=0}^{n-1}[\Phi^{\delta}]^{n-i-1} E_{v_i}^{h,\delta} \sum_{l=v_i+1}^{v_{i+1}-1} \left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta)\right) - p(\theta)\delta\bar{y}_l^{0,h,\delta}\right]^2 I_{\{\xi_l^{0,h,\delta}\notin G_h\}}.$$

Now, using the conditional variance computation (6.12), we see that the contribution to (6.10) due to the reflection steps is

$$O(h)E\left[\left|z^{0,h,\delta}\right|^2(t-\tau_0) - \left|z^{0,h,\delta}\right|^2(t-\tau)\right].$$

By Lemma 5.1, this satisfies (5.8).

Next consider the errors due to the non-reflection steps. Suppose that it has been decided that time does not advance at step n and that $\xi_n^{0,h,\delta} \in G_h$. Then, given this information, the conditional mean of the increment $\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta)$ is $q_n^{h,\delta}(\theta)$, with $p(\theta)\delta\bar{y}_n^{0,h,\delta} = 0$. The conditional variance of $\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta)$ is bounded above by that obtained if each of the two remaining components were randomized separately. So, first consider the term $b(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\Delta t_n^h$ and set the term involving $\mu(\cdot)$ to zero. Suppose (w.l.o.g.) that we have centered the $b(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\Delta t_n^h$ so that they are in $[0,h]$. Then choose h with probability $b(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\Delta t_n^h/h$. This yields a conditional variance of order $O(h)\Delta t_n^h$. Now redefine

$$Q_n^{h,\delta}(\theta) = \sum_{l=v_n+1}^{v_{n+1}-1}\left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta)\right) - b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta)\Delta t_l^h\right] I_{\{\xi_l^{0,h,\delta}\in G_h\}}.$$

As above, the summands are martingale differences and the mean square value of (6.13), with the new definition of $Q_n^{h,\delta}(\theta)$ used, is

$$\begin{aligned} E\sum_{i=0}^{n-1}[\Phi^{\delta}]^{n-i-1} E_{v_i}^{h,\delta} \sum_{l=v_i+1}^{v_{i+1}-1} &\left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta)\right) - b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta)\Delta t_l^h\right]^2 \\ &\times I_{\{\xi_l^{0,h,\delta}\in G_h\}}. \end{aligned} \tag{6.14}$$

Since $E_{v_n}^{h,\delta}[v_{n+1} - v_n] = O(1/h)$, and the expectation (conditioned on $\mathcal{F}_{v_n}^{h,\delta}$) of the typical summand of the inner sum of (6.14) is $O(h)\Delta t_l^h = O(h^3)$, the expectation (conditioned on $\mathcal{F}_{v_n}^{h,\delta}$) of the inner sum is $O(h^2)$. So (6.14) is $O(h^2)$

times the mean number of time updates needed to get to interpolated time t. But this is $O(t/h)$. So the contribution to the left side of (6.10) is zero.

Now consider the component of $q_n^{h,\delta}(\theta)$ defined by (setting $b(\cdot) = 0$)

$$\hat{q}_n^{h,\delta}(\theta) = \frac{[\mu(\theta) - \mu(\theta - \delta)]}{\delta} g(\xi_n^{0,h,\delta}, u_n^{h,\delta}, \theta)\Delta t_n^h.$$

Again, suppose that we have centered so that it is in $[0, h]$. Then select h with probability $\hat{q}_n^{h,\delta}(\theta)/h$. The conditional (on $\mathcal{F}_n^{h,\delta}$) variance of $\xi_{n+1}^{1,h,\delta}(\theta) - \xi_n^{1,h,\delta}(\theta)$ is $O(h)\hat{q}_n^{h,\delta}(\theta) + [\hat{q}_n^{h,\delta}(\theta)]^2 = O(h)\hat{q}_n^{h,\delta}(\theta)$. To evaluate the total effect on the interval $[0, t]$, redefine

$$Q_n^{h,\delta}(\theta) = \sum_{l=v_n+1}^{v_{n+1}-1} \left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta) \right) - \hat{q}_l^{h,\delta}(\theta) \right] I_{\{\xi_l^{0,h,\delta} \in G_h\}}.$$

Then evaluate (6.13) with this new definition of $Q_n^{h,\delta}(\theta)$. Since we set $b(\cdot) = 0$, the summands are martingale differences in that

$$E_{v_n+l}^{h,\delta} \left[\left(\xi_{v_n+l+1}^{1,h,\delta}(\theta) - \xi_{v_n+l}^{1,h,\delta}(\theta) \right) - \hat{q}_{v_n+l}^{h,\delta}(\theta) \right] I_{\{\xi_{v_n+l}^{0,h,\delta} \in G_h\}} I_{\{v_n+l<v_{n+1}\}} (1 - I_{v_n+l}^{h,\delta}) = 0.$$

With the new definition of $Q_n^{h,\delta}(\theta)$, the mean square value of (6.13) is

$$E \sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} E_{v_i}^{h,\delta} \sum_{l=v_i+1}^{v_{i+1}-1} \left[\left(\xi_{l+1}^{1,h,\delta}(\theta) - \xi_l^{1,h,\delta}(\theta) \right) - \hat{q}_l^{h,\delta}(\theta) \right]^2. \tag{6.15}$$

The (conditioned on $\mathcal{F}_{v_n}^{h,\delta}$) expectation of the inner sum in (6.15) is

$$\begin{aligned} O(h^2) &\left[(\mu(\theta) - \mu(\theta - \delta))^2 + (\mu(\theta) - \mu(\theta - \delta)) \right] E_{v_i}^{h,\delta}[v_{i+1} - v_i] \\ &= O(h) \left[(\mu(\theta) - \mu(\theta - \delta))^2 + (\mu(\theta) - \mu(\theta - \delta)) \right]. \end{aligned}$$

Now taking the shift $[\Phi^\delta]^{n-i-1}$ in (6.15) into account and noting that $\mu(\theta) - \mu(\theta - \delta) = 0$ for $\theta < -\tau$, we see that (6.15) is $O(h)$, uniformly in n. Thus (6.10) holds. The verification of (6.11) is straightforward and the details are omitted. ■

A continuous time interpolation. Recall the definition of $\Delta\tau_n^{h,\delta}$ below (5.11). Define $\psi^{0,h,\delta}(t) = \xi_n^{0,h,\delta}$, $\psi^{1,h,\delta}(t, \theta) = \xi_n^{1,h,\delta}(\theta)$, and $\phi^{h,\delta}(t) = \phi_n^{h,\delta}$ for $t \in [\tau_n^{h,\delta}, \tau_{n+1}^{h,\delta})$. Recall the discussion of the boundaries on $\xi_n^{1,h,\delta}(\theta)$ at the beginning of this section. They are ignored in the next theorem but reintroduced in Theorem 6.3.

Theorem 6.2. *Assume* (A2.1)–(A2.5) *and that* $\Delta t^h(x^0, x^1, \alpha) = O(h^2)$. *Then*

$$
\begin{aligned}
\int_0^T \psi^{1,h,\delta}(t,0)dt = &\int_0^T dt \int_{-\tau}^0 d\gamma \int_U b(\psi^{0,\delta}(t+\gamma), \alpha, \gamma) m^{h,\delta,\prime}(d\alpha, t+\gamma) \\
&+ \int_0^T dt \int_{-\tau}^0 p(\gamma) d_\gamma \bar{y}^{0,h,\delta}(\gamma + t) \\
&+ \int_0^T dt \int_{-\tau}^0 d\mu(\gamma) \int_U g(\psi^{h,\delta}(\gamma+t), \alpha, \gamma) m^{h,\delta,\prime}(d\alpha, \gamma + t) + \rho_0^{h,\delta}(T),
\end{aligned}
\tag{6.16}
$$

where

$$
\lim_{h,\delta\to 0} \sup_{u^{h,\delta},\hat{x},\hat{u}} E \sup_{s \le t} |\rho_0^{h,\delta}(s)| = 0. \tag{6.17}
$$

Proof. Continuing to use $v_n = v_n^{h,\delta}$, we can write

$$
\xi_{v_{n+1}}^{1,h,\delta}(\theta) = \xi_{v_n+1}^{1,h,\delta}(\theta - \delta) + B_n^{h,\delta}(\theta) + P_n^{h,\delta}(\theta) + G_n^{h,\delta}(\theta) + R_n^{h,\delta}(\theta). \tag{6.18}
$$

where $R_n^{h,\delta}(\theta)$ was defined in Theorem 6.1 and

$$
B_n^{h,\delta}(\theta) = \sum_{l=v_n+1}^{v_{n+1}-1} b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta) \Delta t_l^h,
$$

$$
P_n^{h,\delta}(\theta) = \sum_{l=v_n+1}^{v_{n+1}-1} p(\theta) \delta \bar{y}_l^{0,h,\delta} = p(\theta) \left[\bar{y}_{v_{n+1}}^{0,h,\delta} - \bar{y}_{v_n}^{0,h,\delta} \right],
$$

$$
G_n^{h,\delta}(\theta) = [\mu(\theta) - \mu(\theta - \delta)] \sum_{l=v_n+1}^{v_{n+1}-1} g(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta) \frac{\Delta t_l^h}{\delta}.
$$

Until further notice ignore the effects of the initial condition (3.3). First consider the contribution of the $b(\cdot)$ terms to $\psi^{1,h,\delta}(t,\theta)$. Their total contribution to $\xi_{v_n+1}^{1,h,\delta}(\theta)$ is, for $n \ge 1$,

$$
\begin{aligned}
\sum_{i=0}^{n-1} [\Phi^\delta]^{n-i-1} B_i^{h,\delta}(\theta) &= \sum_{i=0}^{n-1} B_i^{h,\delta}(\theta - n\delta + i\delta + \delta) \\
&= \sum_{l=0}^{v_n - 1} b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - n\delta + \phi_l^{h,\delta} + \delta) \Delta t_l^h (1 - I_l^{h,\delta}).
\end{aligned}
$$

This is equal to

$$
\sum_{l=0}^{v_n-1} b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - n\delta + \phi_l^{h,\delta} + \delta) \Delta \tau_l^{h,\delta} \tag{6.19}
$$

plus a martingale "error" whose quadratic variation process is $\sum_{l=0}^{n-1} O(\Delta \tau_l^{h,\delta})^2$. Define $\gamma^{h,\delta}(t) = \max\{n : \tau_n^{h,\delta} \le t\}$. The number of time advances that have occurred when interpolated time t is reached is $\phi^{h,\delta}(t)/\delta$. For $\phi^{h,\delta}(t) < t$ the values of $\psi^{1,h,\delta}(t,\theta)$ and $\xi^{1,h,\delta}_{v_{\phi^{h,\delta}(t)/\delta}+1}(\theta)$ differ by the contributions of the iterates in the interval $[v_{\phi^{h,\delta}(t)/\delta}+1,\ \gamma^{h,\delta}(t)-1]$, those that occur before interpolated time t is reached but at or after the last update of the time variable before interpolated time t is reached. By (6.11), the contributions of these terms for $\theta = 0$ is asymptotically negligible. Letting $n = \phi^{h,\delta}(t)/\delta$ and adding these terms, (6.19) becomes

$$\sum_{l=0}^{\gamma^{h,\delta}(t)-1} b(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - \phi^{h,\delta}(t) + \phi_l^{h,\delta} + \delta)\Delta\tau_l^{h,\delta}.$$

By Theorem 5.2 and the continuity of $b(\cdot)$, the last expression equals, modulo an error that satisfies (6.17),

$$\sum_{l=0}^{\gamma^{h,\delta}(t)-1} b\left(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - t + \sum_{k=0}^{l-1}\Delta\tau_k^{h,\delta}\right)\Delta\tau_l^{h,\delta},$$

which for $\theta = 0$ can be written as (modulo an error satisfying (6.17))

$$\begin{aligned}\int_0^t b\left(\psi^{0,h,\delta}(s), u^{h,\delta}(s), -t+s\right)ds &= \int_{\max\{0,t-\tau\}}^t b\left(\psi^{0,h,\delta}(s), u^{h,\delta}(s), -t+s\right)ds \\ &= \int_{\max\{-t,-\tau\}}^0 \int_U b\left(\psi^{0,h,\delta}(\gamma+s), \alpha, \gamma\right) m^{h,\delta,\prime}(d\alpha, s+\gamma)d\gamma,\end{aligned} \tag{6.20}$$

where the last equality uses the change of variable $\gamma = -t+s$, the fact that $b(\cdot,\theta) = 0$ for $\theta < -\tau$, and switches to relaxed control notation. The above mentioned martingale error process has quadratic variation $O(h)$ and satisfies (6.17).

Next, consider the contribution of the $p(\theta)\delta\bar{y}_l^{0,h,\delta}$ terms to $\psi^{1,h,\delta}(t,\theta)$. Following the development for the $b(\cdot)$ terms above, the contribution is

$$\sum_{l=0}^{\gamma^{h,\delta}(t)-1} p(\theta - \phi^{h,\delta}(t) + \phi_l^{h,\delta} + \delta)\delta\bar{y}_l^{0,h,\delta}.$$

By Theorem 5.2, $\phi_l^{h,\delta}$ is asymptotically equivalent to $t_l^{h,\delta}$ and $\tau_l^{h,\delta}$. Thus, by Theorem 5.2 and the continuity of $p(\cdot)$, the right hand side can be written as

$$\sum_{l=0}^{\gamma^{h,\delta}(t)-1} p(\theta - t + \tau_l^{h,\delta})\delta\bar{y}_l^{0,h,\delta}$$

modulo an error that can be written as $\epsilon(h,\delta)\left[|z^{0,h,\delta}|(t) - |z^{0,h,\delta}|(t-\tau-\delta)\right]$, where $\epsilon(h,\delta) \to 0$ as $h,\delta \to 0$. By Lemma 5.1, the error term satisfies (6.17).

Then, for $\theta = 0$, by a change of variable and using the fact that $p(\theta) = 0$ for $\theta < -\tau$, we can write the last expression as

$$\int_0^t p(-t+s)d\bar{y}^{0,h,\delta}(s) = \int_{\max\{-t,-\tau\}}^0 p(\gamma)d_\gamma \bar{y}^{0,h,\delta}(t+\gamma). \tag{6.21}$$

Now consider the contribution of the terms involving $g(\cdot)$ to the integral $\int_0^T \psi^{1,h,\delta}(s,0)ds$. By (6.11), to evaluate this integral, we can suppose that $\psi_l^{1,h,\delta}(0)$ is constant on the intervals $[\tau_{v_n}^{h,\delta}, \tau_{v_{n+1}}^{h,\delta})$ between updates of the time variable. Define $N^{h,\delta}(T) = \phi^{h,\delta}(T)/\delta$. Then, making the piecewise constant approximation, the integral is (modulo an error satisfying (6.17)),

$$\sum_{n=0}^{N^{h,\delta}(T)-1} \xi_{v_n}^{1,h,\delta}(0)\left[\tau_{v_{n+1}}^{h,\delta} - \tau_{v_n}^{h,\delta}\right]. \tag{6.22}$$

Until further notice, consider only the part that is due to the $g(\cdot)$ terms and continue to ignore the effects of the initial condition. Then we have, for $n \geq 1$,

$$\xi_{v_n}^{1,h,\delta}(\theta) = \sum_{i=0}^{n-1} G_i^{h,\delta}(\theta - n\delta + i\delta + \delta).$$

By the definition of $G_i^{h,\delta}(\theta)$, this last expression can be written as

$$\begin{aligned}\sum_{i=0}^{n-1} & \frac{[\mu(\theta - n\delta + i\delta + \delta) - \mu(\theta - n\delta + i\delta)]}{\delta} \\ & \times \sum_{l=v_i+1}^{v_{i+1}-1} g(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - n\delta + i\delta + \delta)\Delta t_l^h.\end{aligned} \tag{6.23}$$

The sum (6.22) will only be changed by a quantity satisfying (6.17) if we replace the inner sum in (6.23) by

$$\hat{G}_i^{h,\delta}(-n+i) = \sum_{l=v_i}^{v_{i+1}-1} g(\xi_l^{0,h,\delta}, u_l^{h,\delta}, \theta - n\delta + i\delta + \delta)\Delta\tau_l^{h,\delta}.$$

Recall that $g(\cdot,\theta) = 0$ and $\mu(\theta) = 0$ for $\theta \leq -\tau$. Let $\theta = 0$. By a change of variable $n - i = q$, the use of $\hat{G}_i^{h,\delta}(\cdot)$ in lieu of $G_i^{h,\delta}(\cdot)$, and a change in the order of summation, we can write (6.22) as

$$\sum_{q=1}^{N^{h,\delta}(t)-1} [\mu(-q\delta + \delta) - \mu(-q\delta)] \sum_{n=q}^{N^{h,\delta}(T)-1} \hat{G}_{n-q}^{h,\delta}(-q)\frac{\left[\tau_{v_{n+1}}^{h,\delta} - \tau_{v_n}^{h,\delta}\right]}{\delta}. \tag{6.24}$$

The fraction on the right can be replaced by unity, only incurring an error that satisfies (6.17). Then we can write the inner sum as (modulo an error satisfying (6.17))

$$\int_0^{T-q\delta} g(\psi^{0,h,\delta}(s), u^{h,\delta}(s), -q\delta + \delta)ds.$$

Next, using this expression and the continuity of $g(\cdot)$, (6.24) can be approximated by

$$\int_{\max\{-T,-\tau\}}^{0} \mu(d\gamma) \int_0^{T+\gamma} g(\psi^{0,h,\delta}(s), u^{h,\delta}(s), \gamma) ds$$

which is equal to

$$\int_{-\tau}^{0} \mu(d\gamma) \int_{-\gamma}^{T} g(\psi^{0,h,\delta}(t+\gamma), u^{h,\delta}(t+\gamma), \gamma) dt. \tag{6.25}$$

By Theorem 6.1, the contribution of the randomization errors $\rho_l^{h,\delta}(\theta)$ to the left side of (6.16) satisfies (6.17).

It can be shown, via analogous computations, that adding the effects of a discretization of the initial condition (3.3) changes the inner integral in (6.25) to $\int_0^T$, and in (6.20) and (6.21) changes the integral to $\int_{-\tau}^0$. The few details are omitted. Finally, making these changes, integrating the resulting (6.20) and (6.21) over $[0, T]$, and writing (6.25) in relaxed control notation, yields the theorem. ∎

Convergence of the numerical algorithm. Recall the representation (6.5). The continuous time interpolation, $M^{h,\delta}(\cdot)$ (with intervals $\Delta\tau_n^{h,\delta}\}$) of the martingale $\sum_{i=0}^{n-1} \beta_i^{0,h,\delta}$ is a martingale with quadratic variation process $\int_0^t a(\psi^{0,h,\delta}(s))ds$ plus an error that goes to zero as $h, \delta \to 0$. The next theorem shows that the optimal values computed by the numerical algorithm converge to the optimal value of the original problem as $h, \delta \to 0$.

Theorem 6.3. *Assume* (A2.1)–(A2.5) *and that* $\Delta t^h(x^0, x^1, \alpha) = O(h^2)$. *Suppose that there is no delayed reflection term and that the boundaries on* $\xi_n^{1,h,\delta}(\theta)$ *are large enough so that they would not be exceeded by* $\chi^1(t, \theta)$. *Then there is a martingale* $w^{h,\delta}(\cdot)$ *with quadratic variation process* It, *that converges weakly to a Wiener process, and for which* (*modulo a term that goes to zero as* $h, \delta \to 0$)

$$M^{h,\delta}(t) = \int_0^t \sigma(\psi^{h,\delta}(s) dw^{h,\delta}(s). \tag{6.26}$$

For any sequence of controls for the chain, the set $(\psi^{0,h,\delta}(\cdot), m^{h,\delta}(\cdot), w^{h,\delta}(\cdot), z^{0,h,\delta}(\cdot))$ (*interpolation intervals* $\Delta\tau_n^{h,\delta}$) *is tight in the Skorohod topology and converges weakly to a solution to* (2.3). *The optimal costs for the chain* $\{\xi_n^{0,h,\delta}, \xi_n^{1,h,\delta}(\theta), \theta \in T^\delta\}$, *and cost function* (6.8) *converge to the optimal cost for original process* (2.3) *and cost function* (2.6) *if the initial conditions are the same.*

Now add the delayed reflection term and recall the discussion on boundaries at the beginning of the section. The sequence $(\xi^{0,h,\delta}(\cdot), m^{h,\delta}(\cdot), w^{h,\delta}(\cdot), z^{0,h,\delta}(\cdot))$ *is still tight and the limits of the optimal costs for the chain are arbitrarily close to that for the original process if the boundaries are large enough.*

Proof. With the preparation in Theorem 6.2 in hand, the proof follows that in [10, Chapter 11] closely and the reader is referred to that reference for more detail. Fix a control sequence. The martingale $M^{h,\delta}(\cdot)$ is tight in the Skorohod topology. Then, since the increments $\beta_n^{0,h,\delta}$ are $O(h)$, any weak-sense limit has continuous paths with probability one. As noted above and in Section 5. the proof [10, Section 10.4.1] implies that there is a martingale $w^{h,\delta}(\cdot)$ satisfying (6.26), modulo an asymptotically negligible error, and that converges weakly to a standard vector-valued Wiener process. The error is due to the $o(\Delta t_n^h)$ term in (6.4). Lemma 5.1 is applicable and implies that $z^{0,h,\delta}(\cdot)$ is tight. Again, since the $\delta z_n^{0,h,\delta}$ are $O(h)$, the tightness implies that all weak-sense limit processes are continuous w.p.1. These facts and the boundedness of $\xi_n^{1,h,\delta}(0)$ implies the tightness of $\psi^{0,h,\delta}(\cdot)$ and the asymptotic continuity of any weak-sense limit. Any sequence $m^{h,\delta}(\cdot)$ of relaxed controls is tight.

Now extract a weakly convergence subsequence with limit denoted by $(x(\cdot), w(\cdot), m(\cdot), z(\cdot))$. Then the proofs in [10, Chapters 10, 11] imply that $(x(\cdot), w(\cdot), m(\cdot), z(\cdot))$ is nonanticipative with respect to $w(\cdot)$, that $m(\cdot)$ is an admissible relaxed control, and that the set satisfies (2.3). Also, the costs for the chain converge to the cost for the limit process. If we let $u^{h,\delta}(\cdot)$ be an optimal control for the chain, then these comments imply that $\liminf_{h,\delta\to 0} V^{h,\delta}(\hat{x},\hat{u}) \geq V(\hat{x},\hat{u})$. The reverse inequality $\limsup_{h,\delta\to 0} V^{h,\delta}(\hat{x},\hat{u}) \leq V(\hat{x},\hat{u})$ is proved just as it was for the no-delay problem in the reference. The presence of delays does not change the structure and the details are omitted. ■

7 Size of the state space for the approximating chain.

The comments concerning dimension and memory below (3.3) all apply to the numerical procedures. Note that the complexity of the computation of $\xi_n^{1,h,\delta}(\theta)$ is not heavily dependent on the dimension of the control variable, or on components of $x(\cdot)$ that do not have delay components. If the control and reflection term are not delayed, then [6] discusses useful numerical procedures, based on the use of (2.2), without the need for auxiliary variables, and they are preferable for such problems. Those procedures do not appear to be useful if the control and/or the reflection processes are also delayed. In particular, one would have to keep track of the values of the control the reflection terms over the delay intervals and approximate them by finite-valued discrete-time processes that do not lose too much information. These approximations become part of the state space of the approximating chain, usually making its size much too large. For such problems the approach of this paper is quite promising. The size of the state space is the product of what is needed for $\xi_n^{0,h,\delta}$ and $\xi_n^{1,h,\delta}(\theta)$, where θ takes τ/δ values.

Consider a one-dimensional problem where $g(\cdot) = p(\cdot) = 0$ and let the discretization level for the $\xi_n^{1,h,\delta}(\theta)$ be h. Then there is a $K < \infty$ such that $|\chi^1(t,\theta)| \leq K(\tau+\theta), -\tau \leq \theta \leq 0$. Without loss of generality, suppose that $x(t)$ has been centered so that it lies in an interval $[0, B_0]$ for some $B_0 < \infty$. The state space for $\xi_n^{0,h,\delta}$ has $[B_0/h+3]$ points. There are τ/δ values for θ. Thus, if we bound $\xi_n^{1,h,\delta}(\theta)$ by $K(\tau+\theta)$, the maximum number of points (which includes the reflecting states) is

$$[B_0/h+3]\frac{K\tau}{h}\frac{K(\tau-\delta)}{h}\cdots\frac{K\delta}{h}.$$

While large, this is much better than a direct procedure (or the procedure in [6]) when there are delays in the control. We will now improve it further.

Let $h=\delta$ and suppose that $b(\cdot)$ is Lipschitz continuous in θ, uniformly in the other variables. Then

$$\xi_n^{1,h,\delta}(\theta)-\xi_n^{1,h,\delta}(\theta-\delta)=O(\delta)+\text{ noise,} \tag{7.1}$$

where the noise is due to the randomization and has variance (Theorem 6.1) $O(h\Delta t_n^h)=O(h^3)$. Since $\xi_n^{1,h,\delta}(-\tau)=0$, the state space can consist of the differences $\xi_n^{1,h,\delta}(\theta)-\xi_n^{1,h,\delta}(\theta-\delta)$, $-\tau+\delta\leq\theta\leq 0$. Recall the discussion in Theorem 6.1 concerning the randomization noise. In the example there, when doing the randomization to attain the desired conditional mean, the value h was selected with a probability of order $O(h)$, and the value zero with a high probability. This implies that, with a high probability, the number of selections of the value h between advances of the time variable is bounded by some small number. Thus, suppose that we can (with a high enough probability) bound the differences in (7.1) by $H\delta$, for some integer H. Then the size of the state space is $[B_0/h+3][H]^{\tau/h}$ a considerable improvement. More care is needed if there is a there is a $g(\cdot)\mu$ or δy term. But there will still be an improvement. Clearly, much more work is needed on the algorithms and state space representations.

References

[1] E. Altman and H. J. Kushner. Heavy traffic analysis of AIMD models for congestion control. In S. Raghavan and G. Anandalingham, editors, *Telecommunications Planning: Innovations in Pricing, Network Design and Management*. Klewer, Amsterdam, 2005. Selected papers from the 2004 INFORMS telecommunications conf,.

[2] R. F. Curtain and H. J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory*. Springer-Verlag, Berlin and New York, 1995.

[3] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics Stochastic Rep.*, 35:31–62, 1991.

[4] P. Dupuis and H. Ishii. SDE's with oblique reflection on nonsmooth domains. *Ann. Probability*, 21:554–580, 1993.

[5] J.M. Harrison and R.J. Williams. Brownian models of open queueing networks with homogeneous customer populations. *Stochastics Stochastics Rep.*, 22:77–115, 1987.

[6] H.J.Kushner. Numerical approximations for nonlinear stochastic systems with delays. *Stochastic and Stochastic Reports*, 77:211–240, 2005.

[7] A. Ichikawa. Quadratic control of evolution equations with delays in control. *SIAM J. Control and Optim.*, 20:645–668, 1982.

[8] H.J. Kushner. *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*. Springer-Verlag, Berlin and New York, 2001.

[9] H.J. Kushner and D. Barnea. On the control of linear functional-differential equations with quadratic cost. *SIAM J. Control Optim.*, 8:257–272, 1970.

[10] H.J. Kushner and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, Berlin and New York, 1992. Second edition, 2001.

[11] A. Manitius and T. Tran. Numerical approximations for hereditary systems with input and output delays: Convergence results abd convergence rates. *SIAM J. Control and Optimiz.*, 32:1332–1363, 1994.

[12] M.I. Reiman and R.J. Williams. A boundary property of semimartingale reflecting Brownian motions. *Prob. Theory Rel. Fields*, 77:87–97, 1988.

[13] R. B. Vinter and R. H. Kwong. The infinite time quadratic control problem for linear systems with state and control delays: An evolution equation approach. *SIAM J. Control and Optim.*, 19:139–153, 1981.